







FYLVNSYDDVISFEKNPTEYINVDDSPKKLEQINNFDYNNNNNNNPIENEGFCLV
SUNDGQFVKGDKSCCIYDKNYIYIFQNKRMQTFINNDYNNQNFIDILNOLKN
KNDLNTLYLYEILCHALYELGKHRIYPGTLIQLSAIKFLALFYKKNKLFSE
PLRKYEQNYQTFLLDCDIPHIQKLOKOKFNKNSTKNQNTINTPVNNWTKRDIKYHK
YINLYDKILYSOTIITNKTKQKKIKYT"
23110..24606
/gene="PFC0340w, MAL3P3.4"
23110..24606
/gene="PFC0340w, MAL3P3.4"
/notes="PFC0340w, (MAL3P3.4), putative DNA polymerase delta
small subunit, len: 499 aa; Similarity to DNA polymerase
delta, small subunits: A.thaliana DNA polymerase delta,
small subunit (F.R.048520) BLAST Score: 345, sum P(3) =
1.3e-58; 35% in 208 aa overlap, predicted using hexexon"
/codon\_start=1
/product="putative DNA polymerase delta small subunit"
/protein\_id="CAB11105.1"
/db\_xref="GI:3649756"
/db\_xref="SPTREMBL:O77321"
/translation="MDEVVTNOEVEEIIKHNNKKNKEGHQNKIKKKHFEYVNLSE
RFIITCPYTOQFSHIYSLRIKIMKTLMTIEDIVKKEGNEKKKNINNDINND
NINNDINNDINNDINNDINNDINNDINNDINNDINNDINNDINNDINNDINND
KMLRPSILNEYSILNNTINVDHEDILFLEDEARLKLEGINSDVFYIGTLV
IVKGIGMSGSLYVEEIIYSIKPLDIPNCISNDDEKFLVSGLYINELKSNINSL
LBNFLRLHGDTYLSENLRILIGNSLSDNDEKDMNTIDVFLSLSCSIHIDLP
GDKPDSNLPQPPNFFKARNFSQCVTPNPFYSIDNINICCMGSEFVNITS
YSKNNKMLALIDIASRILSPSTDLGCTPFTKNDPFLDNDNTYPIHFINGNSKL
EYQVMQDHKQLPLVLPNPHISPKAFILNKLNEYKILTFQIDEDK"
25880..31015
/gene="PFC0345w, MAL3P3.5"
25880..31015
/gene="PFC0345w, MAL3P3.5"
/notes="PFC0345w, (MAL3P3.5), Hypothetical protein, len:
1712 aa"
/codon\_start=1
/product="hypothetical protein, PFC0345w"
/protein\_id="CAB11106.1"

Query Match 7.4%; Score 52; DB 3; Length 116696;
Best Local Similarity 44.6%; Pred. No. 1.6;
Matches 247; Conservative 0; Mismatches 305; Indels 2; Gaps 1;

QY 111 AAAACATTCTACTAAATCGAATTAATTTGAGCTGGCTTGATTAATGGTAGTGG 170
Db 103039 AATAATAATATTAAGAGATGAGCATAAAGAGATAACAATGAAGAATGATAAGATG 103098

QY 171 CTCGGAGGTGATATTTATCCTTGATAGGAACACGCTCTTGAAGCTGTAAGTTATTCA 230
Db 103099 AGTGAATGATATGTTATTAACAATAAAGAGATATGATCATATAATAATAATAT 103158

QY 231 ATCGCTTGTTTCTATACAAAAACATGTTTGAAGAAAGATGAAGACATTTGCTTAAA 290
Db 103159 AACAAATATGATGATGATAATAATAATAACAACATGATGATGATAATAATAAT 103218

QY 291 CCCTTTAATCGCGCTGAGATAAATGACTTTTGGTCTCGAAGAAAGACTTACGCAACT 350
Db 103219 AACACAATATGATGATGATGATAATAATAACAACATATGATGATGATAATAAAC 103278

QY 351 CTCCTAATCTCTTAATCTGAATPATTATCTTTTCCCTTACATGGAATAAATCTGCTTGA 410
Db 103279 AACACAATAATGATGATGATAATAATAAGATCCCAATAGATGGAGAACTCGTCT 103338

QY 411 TATATTT--CCAAATTAACACAATAAAATCAACAAAATTTCTCTTGAAAGTTCGGAGAAA 468
Db 103339 TATATTGATGAAATAAATCAAAAAGAAATAAATATATATTTCTTAAATATATAGATAAATA 103398

QY 469 ACTCAATCTCCCAATTTGATATTTGGATATTATCGATGGATTTCTTCTTTTAATA 528
Db 103399 ATTGAATGTTGAACGAATGTTATTATGAAAAATAACGTTTGTGATGGTATTATTAT 103458

QY 529 AACATCGCGCTATAAAGAAAAATGCCACTTTGAATAATCAAGAAATCAAAATATTCAAAAT 588
Db 103459 TTAATTAGACAGAAAATGGTTTGAATATTTGTCAGAAAGTTAATAACATAATATAA 103518

QY 589 GAAATAATGAAGCCACTAACTCAACAAATCCAGAAATCAGTGGGTAATACTTAACACAG 648
Db 103519 AACAAAATGAAGCAAGAAAGTTTTATATAAAAAAACAATAAATAATACACAG 103578

QY 649 AACCAAAACACAAA 662
Db 103579 AACAGAACAAGAA 103592

RESULT 4
AC073883/c 211376 bp DNA linear ROD 12-FEB-2002
LOCUS Mus musculus chromosome 3 clone rp23-71el3, complete sequence.
DEFINITION AC073883
ACCESSION AC073883
VERSION AC073883.32 GI:18390271
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 211376)
AUTHORS Hu, X., Swank, R. and Roe, B.A.
TITLE Mus musculus Chromosome 3 BAC Clone rp23-71el3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 211376)
AUTHORS Hu, X., Swank, R. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 211376)
AUTHORS Hu, X., Swank, R. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (12-FEB-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Jan 29, 2002 this sequence version replaced gi:18201802.
-----
Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES
source Location/Qualifiers
1..211376
/organism="Mus musculus"
/db\_xref="taxon:10090"
/chromosome="3"
/clone="rp23-71el3"
/clone\_lib="RPI Mouse BAC Library 23"
BASE COUNT 60567 a 41306 c 42611 g 66892 t
ORIGIN

Query Match 7.4%; Score 51.6; DB 10; Length 211376;
Best Local Similarity 54.1%; Pred. No. 1.7;
Matches 105; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 423 AACACAAATAAATCAACAAAATTTCTTGAAAGTTCGGGAGAAACTCAATCTCCACA 482
Db 52670 AAAACAAAACAAAACACACCTTTTCTCCANTGGGGGAAAAAACAACACACCC 52611

QY 483 ATTGGATATTGGAGPATTATCGATGGATTTTCTTCTTTTAAATACATCGCGCTATA 542
Db 52610 CCAGGTTGTGCTGTPAGAATGCTTGAAATGCTACTTGCTAAATTTAGCAAGTGTGTA 52551

QY 543 AAAGAAATGCCACTTGAATAATCAAGAAATCAAAATATTCAAAATGAATAATGAAGCC 602
Db 52550 AAACCAACAAAAGAACACCAACCAAAAAAAGAAAAAGAAAAACCAACACA 52491

QY 603 ACTAACTCAACA 616
Db 52490 AACAAAACACAA 52477



## RESULT 5

PFMAL13P9 78578 bp DNA linear HTG 11-AUG-1999  
LOCUS Plasmodium falciparum chromosome 13 strain 3D7, \*\*\* SEQUENCING IN  
DEFINITION PROGRESS \*\*\*, in unordered pieces.

ACCESSION AL096783.2 GI:5731889

VERSION HTG; HTGS\_PHASE1.

KEYWORDS malaria parasite P. falciparum.

SOURCE Plasmodium falciparum.

ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 78578)

AUTHORS Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M.

and Barrell, B.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-1999) P. falciparum Genome Sequencing Consortium,

The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge

CB10 1SA, UK

COMMENT On Aug 12, 1999 this sequence version replaced gi:5420293.

For more information about this sequence or the Malaria Project,

see <http://www.sanger.ac.uk/projects/P.falciparum>. IMPORTANT: This

sequence is unfinished and does not necessarily represent the

correct sequence. Work on the sequence is in progress and the

release of this data is based on the understanding that the

sequence may change as work continues. The sequence may be

contaminated with foreign sequence from E.coli, yeast, vector,

phage etc.

Order of segments is not known; 800 n's separate segments.

\* NOTE: This is a 'working draft' sequence.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

## FEATURES

source Location/Qualifiers

1..78578

/organism="Plasmodium falciparum"

/strain="3D7"

/db\_xref="taxon:5833"

/chromosome="13"

BASE COUNT 26762 a 9436 c 7899 g 27281 t 7200 others

## ORIGIN

Query Match 7.3%; Score 51; DB 2; Length 78578;  
Best Local Similarity 51.6%; Pred. No. 2.7;  
Matches 142; Conservative 0; Mismatches 130; Indels 3; Gaps 1;

Qy 363 AATTCTGAATATATCTTTCTTACATGGAATAACTGCATCTTGGATATATCCAAATT 422

Db 30299 AATTAAAAAATATATATTTTCTGATTTTACACGGATTTTTCAGAATATATATAAT 30358

Qy 423 AACACAATAAATAACACAAAATTTCTTGTGAAAGTTCCGGAGAAAACATCAATCTCCACA 482

Db 30359 ATGTTAGAAAAAAGAAAGGATTTTTCGATTTTAAAGAAATAAATAAATAAATAA 30418

Qy 483 ATTGATATGGAGTATATCGATGGATTTTCTTCTTTTAAATAACATCGCGCTATA 542

Db 30419 TTTAATCATATAATATATATATTTTCTGATTTTAAATAAATCTGCATCTTCTA 30475

Qy 543 AAGAAAATGCCACTTGAATAATCAAGAAATCAAAATATTCAAAATGAATAATGAAGCC 602

Db 30476 AAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 30535

Qy 603 ACTAACTCAACAAATCCAGAAATCAGGTGGGAATA 637

Db 30536 ATTGAAAACATTTATCTGGATTTCTGCTTTCAGTA 30570

## RESULT 6

AX346425/c 6928 bp DNA linear PAT 01-FEB-2002  
LOCUS Bombyx mandarina mitochondrial DNA, complete genome.

DEFINITION Bombyx mandarina fifth instar larva silkland mitochondrion DNA.

ACCESSION AX346425

VERSION AX346425.1 GI:18494311

## KEYWORDS

SOURCE synthetic construct.

ORGANISM Bombyx mandarina

REFERENCE 1 (sites)

AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.

JOURNAL Diagnosis of diseases associated with the immune system

Patent: WO 0200928-A 1496 03-JAN-2002;

Epigenomics AG (DE)

FEATURES Location/Qualifiers

source 1..6928

/organism="synthetic construct"

/db\_xref="taxon:32630"

/notes="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 1852 a 151 c 1572 g 3353 t

ORIGIN

Query Match 7.3%; Score 50.8; DB 6; Length 6928;

Best Local Similarity 47.6%; Pred. No. 4.8;

Matches 180; Conservative 0; Mismatches 197; Indels 1; Gaps 1;

Qy 304 CTGAGAATAAATGACCTTTTGGTTCTGAGAAAGACTTAGCGACAACCTCTCCAATCTTCTA 363

Db 4902 CTAATCATCACTATTTTAAAAAATTAACAACTAAATTTACTACCTCTTAATTTACTA 4843

Qy 364 ATTCGATATATATCTTTTCTTACATGGAATAACTGCATTTCTGGATATATTTCAATTA 423

Db 4842 ATCTATACTAACATCATTAATACCCAAACAAAACATTTTACTTTTATMAAACCTAAAT 4783

Qy 424 ACACAAATAAATCAACAAAATTTCTTGAAGTTCGGGAGAAAACCTCAATCTCCACAA 483

Db 4782 AAACATAAATACCCGATATAAATATCTATATACATAAATAAATAAATAAATAAATA 4723

Qy 484 TTGATATTTGGATATATATCATGATGATTTTCTTCTTTTAAATTAACATCGCGCTATAA 543

Db 4722 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4664

Qy 544 AAGAAAATGGCCACTTGAATAATCAAGAAATCAAAATATTTCAAAATGAATTAATGAAGCA 603

Db 4663 AAAAAATCACACAAACTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 4604

Qy 604 CTAACACTCAACAAATCCAGATCAGTGGGGAATAACTTAACCTAGAACCAACCAACAAAG 663

Db 4603 TACTAACTACAAATAACTTCAACACCATATTTATACCTAAAAAATAAATAAATAAATA 4544

Qy 664 CCTCATCCAGTTGTTAGA 681

Db 4543 ACCACACAAATTTCTATA 4526

## RESULT 7

AB070263/c

LOCUS Bombyx mandarina mitochondrial DNA, complete genome.

DEFINITION Bombyx mandarina fifth instar larva silkland mitochondrion DNA.

ACCESSION AB070263

VERSION AB070263.1 GI:18640048

KEYWORDS

SOURCE Bombyx mandarina

ORGANISM Bombyx mandarina

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;

Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (sites)

AUTHORS Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K., Nho, S. K., Zhong, B.

and Banno, Y.

TITLE Not trivial level of sequence divergence and sequence

rearrangements of mitochondrial genome between the wild mulberry

silkmoth, Bombyx mandarina and its close relative, the domesticated

silkmoth Bombyx mori

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 15928)

AUTHORS Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K. and Banno, Y.

TITLE Direct Submission



```
gene VKTNTFYGLRSSNK"
5391..6548
/genes="cob"
CDS 5391..6548
/genes="cob"
/codon_start=1
/transl_table=5
/evidence-not_experimental
/product="cytochrome b"
/protein_id="BAB84646.1"
/db_xref="GI:18640054"
/translators="MEFESIRKTHPIIKIINSLIDLPSPNISLSSWNGSLLALCLM
IQITGLFLMYTANEMAFYSVNICRVNMGWIRLHANGASFFFCIYLHGR
GIYFSNLYWGLIMFLMLAFYVLPWGMGFWGATVITNLUSAIPYLG
MLVNWGGFADNATLTRYTFHLLPFIILMTLHLLFLHGTGNSNPLGNSLD
KIPAFPFYFDKIDILIKMLTLTLINPYLIGDPNFTPANPLVTPVHIQPEWYF
LFAYILRSPNKLGGVIALIMSILILIPFTFKKIGIQIOPVPINOIIFWMLMIT
ILLTWIGARVENPIYTIGQLLTIMYFLYFLNPIIIGMYWDKILFNKKNN"
6548..6613
tRNA /genes="cox3"
/notes="codon recognized: UCA; codon recognized:UCN"
/product="trna-Ser"
/anticodon="(pos:6578..6580,aa:Ser)
/evidence-not_experimental
complement(6639..7583)
gene="nad1"
complement(6639..7583)
CDS /genes="nad1"
/codon_start=1
/transl_table=5
/evidence-not_experimental
/product="NADH dehydrogenase subunit 1"
/protein_id="BAB84647.1"

Query Match 7.38; Score 50.8; DB 3; Length 15928;
Best Local Similarity 48.08; Pred. No. 4.1; Indels 0; Gaps 0;
Matches 145; Conservative 0; Mismatches 157;

Oy 357 TCCTTAAATTCGAATATATCTTTTCCCTTACATGGAATACTGCATCTCTGGATATATT 416
Dy 10433 TTTTATTTTAAATGAATATATAGCTTGAATATGATTTATATTTAACTTAATATTATA 10374

Oy 417 CCATTAACACAATAAATCAACAATAATTTCTCTGGAAGTTCCGGAGAAATCAATC 476
Dy 10373 AGAAAAAATAAAGAGAATAATATAAATTTTAAATTTATATTAATCAATCAATA 10314

Oy 477 TCCACAATTTGGATATTGAGATTTATCGATGATTTCTCTCTCTTTAAATTAACATCGC 536
Dy 10313 TTTAAATTTAAATTTAAATTAATCAATTAATTTATTTATTTATTTATTTATTTATTAAT 10254

Oy 537 CCTATAAAGAAATGCGCACTTGAATTAATCAAGAATCAAAATATTTCAAAATGAAATAT 596
Dy 10253 TTGATTTATTTTAAATAAATAATATTAATTTAAATTTAAATTTAAATTAATTAATCAAA 10194

Oy 597 GAAGCCACTAACTCAACAATTCAGATCAGGTGGGAATTAACTTAACTCAGAACCAAAA 656
Dy 10193 TTAATATTTTAAATTTAAATTAATTAATCAATCAATTAATTTATTTATTTATTTATTAAT 10134

Oy 657 CA 658
Dy 10133 CA 10132
```

```
RESULT 8
AC068138
LOCUS AC068138 168698 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-813K12 from 2, complete sequence.
ACCESSION AC068138
VERSION AC068138.6 GI:15638958
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.  
1 (bases 1 to 168698)  
Sulston,J.E. and Waterston,R.  
Toward a complete human genome sequence  
Genome Res. 8 (11), 1097-1108 (1998)  
99063792  
2 (bases 1 to 168698)  
Isak,A., Kozlowicz,A., Doebber,A. and Dixon,R.  
The sequence of Homo sapiens BAC clone RP11-813K12  
Unpublished (2001)  
3 (bases 1 to 168698)  
Waterston,R.H.  
Direct Submission  
Submitted (28-APR-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 168698)  
Waterston,R.H.  
Direct Submission  
Submitted (18-SEP-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 168698)  
Waterston,R.  
Direct Submission  
Submitted (09-JAN-2002) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Sep 18, 2001 this sequence version replaced gi:15148349.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc  
Contact: saplensew@wustl.edu  
----- Summary Statistics  
-----  
Center project name: H\_NH0813K12  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:  
The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatenoh,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu)  
VECTOR: pBAC3.6  
NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-314B1, 2000 bp overlap; the clone sequenced to the right is RP11-86O17. Actual start of this clone is at base position 161288 of RP11-314B1; actual end is at base position 168698 of RP11-813K12.

Data from AC069383 was used to finish this clone, AC068138.An



Db	97691	ATATATAAATATATTAAATATATATAATATATATTTAAATATATATATATTTAAATAT	-ATAA	97748
Qy	375	TATCTTTTCCTTACATGGAATAACTGCATTCTTGATATATATCCAAATTAACACAATAAAA	434	
Db	97749	TATATATTTAAATATATAATATATTTAAATATATAATATATATATTTAAATGTATATATTTA	97808	
Qy	435	ATCAACAAAATTTCTCTTTGAAAGTTCCGGAGAAACATCAATCTCCACAAATGGATATGG	494	
Db	97809	ATATATATAATATATTTAAATGTATATAATATATTTAAATATATAATATATATTTTGTATATAA	97868	
Qy	495	AGTATTATCGATGATTTCTCTCTCTTTAAATTAACCAT	533	
Db	97869	TATATTTAAATATATATTTGATGTATAATATATTTAAATAT	97907	
RESULT 9				
CNS06EHA				
LOCUS				
DEFINITION		CNS06EHA 263 bp DNA linear STS 10-JAN-2001		
		T3 end of clone AR0AA021E1I of library AROAA from strain CHS 732 o		
ACCESSION		Zygosaaccharomyces rouxii, sequence tagged site.		
VERSION		AL395108		
KEYWORDS		AL395108.1 GI:12146479		
SOURCE		STS.		
ORGANISM		Zygosaaccharomyces rouxii.		
REFERENCE		Zygosaaccharomyces rouxii.		
AUTHORS		Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
		Saccharomycetales; Saccharomycetaceae; Zygosaaccharomyces.		
TITLE		1 (bases 1 to 263)		
JOURNAL		de Montigny,J., Straub,M., Potier,S., Tekala,F., Dujon,B.,		
PUBMED		Wincker,P., Artiguenave,F. and Souciet,J.		
REFERENCE		Genomic Exploration of the Hemiascomycetous Yeasts: 8.		
AUTHORS		Zygosaaccharomyces rouxii(1)		
		FEBS Lett. 487 (1), 52-55 (2000)		
TITLE		11152883		
JOURNAL		2 (bases 1 to 263)		
PUBMED		Souciet,J.L., Aligte,M., Artiguenave,F., Blandin,G.,		
REFERENCE		Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,		
AUTHORS		de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,		
		Malpertuy,A., Neuveglise,C., Ozler-Kalogeropoulos,O., Potier,S.,		
		Saurin,W., Tekala,F., Toffano-Nloche,C., Wesolowski-Louvel,M.,		
		Wincker,P. and Weissenbach,J.		
TITLE		Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of		
JOURNAL		yeast species for molecular evolution studies(1)		
PUBMED		FEBS Lett. 487 (1), 3-12 (2000)		
REFERENCE		11152876		
AUTHORS		3 (bases 1 to 263)		
TITLE		Genoscope.		
JOURNAL		Direct Submission		
PUBMED		Submitted (06-SEP-2000) Genoscope - Centre National de Sequençage,		
REFERENCE		2 rue Gaston Creneauux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :		
AUTHORS		seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
TITLE		This STS is part of a random genomic sequencing program of thirteen		
JOURNAL		yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces		
PUBMED		exiguus, Saccharomyces servazzii, Zygosaaccharomyces rouxii,		
REFERENCE		Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces		
AUTHORS		lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia		
TITLE		angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,		
JOURNAL		Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to		
PUBMED		5 kb were prepared and both extremities were sequenced. See		
REFERENCE		keywords for description of this sequence and for the sequence of		
AUTHORS		the other extremity of this insert.		
TITLE		Location/Qualifiers		
JOURNAL		1 . 263		
PUBMED		/organism="Zygosaaccharomyces rouxii"		
REFERENCE		/strain="CHS 732"		
AUTHORS		/db_xref="taxon:4956"		
TITLE		/clone="AR0AA021E1I"		
JOURNAL		/clone_lib="AROAA"		
PUBMED		/notes="end : T3"		
REFERENCE		1 . 263		
AUTHORS		180 a . 16 c	17 g	23 t
TITLE		27 others		
JOURNAL				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
PUB				







Db	301955	TATAAATTA	CTTCTATATAAATAAATAATATATTTCTTCACAAATTTTACATAATTAATCTTTACATC	301896
Qy	371	ATATTATCTTTTCCATTCATCATGGAATAA	CTGCAATCTCTTGGATATATATTCCTCAATTTTACATAATTAATCTTTACATC	430
Db	301895	ATAAAAAAATTAATTTATTA	TACCAATTAATATATATTAACATCTTTTATAACTCTTTTATAACTATAATAAT	301836
Qy	431	TAAAAATCAACAAAAATTTCTCTTGAAAGTCCGGAGAAAACTCAATCTCCACAATTTGGATA	490	
Db	301835	TTTAAAAAAAATCTTTAATAATTTTCAACACATATAAAAAA	AAAAAATTAATTTAATTTAATTA	301777
Qy	491	TGGAGTATTATCGATGATTTTCTTCTTTTAAATTAACATCGGCCTATAAAGAAAA	550	
Db	301776	TTTAATTAATAATTAATTTATTTATTTCTTCTTCATATAAAAAAT---	ATAATAAAACTAAA	301720
Qy	551	TGGCCACTTGAATTAATCAAGAATCAAAATATTTCAAAATGAATAATGAAGCCACTAAACT	610	
Db	301719	CGCAACTAAATAAAATTTAAAAAATCACTAAACATAATACTCATACTATAATCC	301660	
Qy	611	CAACAAATCCAGAATCAGTGGGAATAACITTAACCTCAGAACCAAAACACAAAGCC	665	
Db	301659	CAACACTTTAAAAAACTAAAAAGAAATAAATCATCTAAAAATCAATAATTTAAAAACC	301605	
RESULT 13				
AX344564/c				
LOCUS	AX344564	349980 bp	DNA	linear
DEFINITION	Sequence 15 from Patent WO0200932.			PAT 01-FEB-2002
ACCESSION	AX344564			
VERSION	AX344564.1	GI:18492450		
KEYWORDS	synthetic construct.			
SOURCE	synthetic construct			
ORGANISM	artificial sequence.			
REFERENCE	1 (sites)			
AUTHORS	Olek.A., Piepenbrock,C. and Berlin,K.			
TITLE	Diagnosis of known genetic parameters within the mbc			
JOURNAL	Patent: WO 020932-A 15 03-JAN-2002;			
FEATURES	Epigenomics AG (DE)			
source	Location/Qualifiers			
	1..349980			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3.673778 <223>-split as follows.-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 649.980-seq 03 600.001 949.980-seq 04 900.001 1.249.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 1.849.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <223>-Original length of seq 2: 3.673778 <223>-split as follows.-seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16 600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"			
BASE COUNT	99616 a 3868 g 72947 g 173548 t		1 others	
ORIGIN				
Query Match	7.2%;	Score 50.2;	DB 6;	Length 349980;
Best Local Similarity	50.1%;	Pred. No. 2.8;		
Matches	178;	Conservative 0;	Mismatches 173;	Indels 4;
Gaps	2;			
Qy	311	TAAAAATGACATTTTGGTCTCAGAAAAGACTTATAGGCACAACTCTCCAACTCTTCTAATCTGA	370	
Db	1955	TATAATTA	CTTCTATAATAATAATTAATTTCTTCACAAATTTTACATAATTAATCTTTACATC	1896
Qy	371	ATATTA	CTTTTCTTACATGGAATAACTGCATCTTTGGATATATTTCCAAATTAATCAACAAA	430

Db	1895	ATAAAAAAATAATTTATTATACAAATTAATAATTTATATATTAACTCTTTTATAAACAATAATAAT	1836
Qy	431	TAATAATCAACAAAAATTTCTCTTTGAAGTTCCGGAGAAAACTCAATCTCCACAATTTGGATA	490
Db	1835	TTTTAAAAAATCTTTAAATATTTTCAACATATAAAAAAATAAAT-AAATTTAAATTA	1777
Qy	491	TTGGAGTATTATCGATGGATTTCTCTCTTTTAATTAACATGCGCTATAAAGAAGAAA	550
Db	1776	TTTAATTAATAATTTATTTATTTCTCTTCATATAAAT---ATAATAAACTAAA	1720
Qy	551	TGGCCACTTTGAATATCAAGAAATCAAAATATTTCAAAATGAATATGAAGCCACTAAACT	610
Db	1719	CGCAACTAATAAATTTAAAAAATCAACTAAACATATAACTCATACCTATAATCC	1660
Qy	611	CAACAAATCCGAATCAGTGGGATCACTTAACCTCAGAACCAACACAAAGCC	665
Db	1659	CAACACTTAAAAAATCAAAAGAAATATCATCTAAAAATCAATAATTTAAAAACC	1605
RESULT 14			
AX281283/c			
LOCUS	AX281283	Sequence 25 from Patent WO0177164.	PAT 02-NOV-2001
DEFINITION	AX281283	11670 bp DNA linear	
ACCESSION	AX281283		
VERSION	AX281283.1	GI:16608538	
KEYWORDS			
SOURCE		synthetic construct.	
ORGANISM		synthetic construct	
REFERENCE		artificial sequence.	
AUTHORS		1 (sites)	
TITLE		Olek.A., Piepenbrock,C. and Berlin,K.	
JOURNAL		Diagnosis of diseases associated with apoptosis	
		Patent: WO 0177164-A 25 18-OCT-2001;	
		Epigenomics AG (DE)	
FEATURES		Location/Qualifiers	
source		1..11670	
		/organism="synthetic construct"	
		/db_xref="taxon:32630"	
		/note="chemically treated genomic DNA (Homo sapiens)"	
BASE COUNT	3287 a	238 c	2377 g
ORIGIN		5768 t	
Query Match	7.1%;	Score 49.6;	DB 6;
Best Local Similarity	48.9;	Pred. No. 7.1;	Length 11670;
Matches	133;	Conservative	0; Mismatches 139; Indels 0; Gaps 0;
Qy	395	TAACGTGATCTTGGATATATCCAAATTAACAAATAAAATCAACAAAATTTCTCTTGA	454
Db	11107	TAATTTATTTTCATCATATACATATACGCTTCTCATATAACTCGACATATATTTTCAAACA	11048
Qy	455	AAGTTCGGAGAAACCAATCTCCAAATTTGGATATTGGAGTATTATCGATGGATTTTC	514
Db	11047	CTAACTTAATTAATACCCCTTTATACAACTACGTATTTTAACATATACCATAACTTTA	10988
Qy	515	TTCTTCTTTAATTAACATGCGCTATATAAAGAAAAATGGCCACTTGAATATCAAGAATC	574
Db	10987	ATTTTAAATTTTCCAAAAATCAGTAACCAATATCTCTCATTTTAAAAATAAAAAATC	10928
Qy	575	AAAATATCAAAATGAATATGAAGCCACTTAAACTCAACAAAATCCAGATCAGGTGGGA	634
Db	10927	TAAAAATACTCCCTCTAATAATAAATAAAAAATAAAAAATTAATAATAAAAAATCAA	10868
Qy	635	ATAACTTAACCTCAGAACCAACCAACAAAGCCT	666
Db	10867	CTAATTAACAATACACAAACAAATAAAAACT	10836
RESULT 15			
AX344573/c			
LOCUS	AX344573	Sequence 24 from Patent WO0200932.	PAT 01-FEB-2002
DEFINITION	AX344573	34980 bp DNA linear	
ACCESSION	AX344573		
VERSION	AX344573.1	GI:18492459	

Search completed: November 2, 2002, 01:41:03  
Job time : 2972.9 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:12:43 ; Search time 153.968 Seconds  
(without alignments)  
7805.766 Million cell updates/sec

Title: us-09-881-556-3

Perfect score: 700

Sequence: 1 gaattggcagcagattttt.....accgcactctacagaaaagc 700

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	50.8	7.3	6928	24	ABL33523 Human immune syste
C 2	49.4	7.1	5814	24	ABL33561 Human immune syste
C 3	48.8	7.0	17280	22	AAS46771 Tumour suppressor
C 4	48.4	6.9	7040	22	AAS46459 Tumour suppressor
C 5	48.2	6.9	9817	24	ABL33368 Human immune syste
C 6	48.2	6.9	17389	24	ABL33414 Human immune syste
C 7	48	6.9	10528	24	ABL32335 Human immune syste
C 8	47.6	6.8	605	17	AT31530 Human 3' apolipop
C 9	47.6	6.8	10039	24	ABL34043 Human immune syste

C 10	47.2	6.7	5501	24	AAS61244 Human gene regulat
C 11	47.2	6.7	9110	24	ABL34428 Human immune syste
C 12	47	6.7	6123	24	ABL32820 Human immune syste
C 13	47	6.7	8392	24	ABL33490 Human immune syste
C 14	47	6.7	17893	24	ABL33365 Human immune syste
C 15	47	6.7	40862	24	ABL34072 Human immune syste
C 16	46.6	6.7	399	22	ABL19737 Human breast cance
C 17	46.6	6.7	19734	24	ABL33933 Human immune syste
C 18	46.4	6.6	5997	24	ABL33624 Human immune syste
C 19	46.2	6.6	10957	24	ABL33110 Human immune syste
C 20	46.2	6.6	15674	24	ABL32363 Human immune syste
C 21	46.2	6.6	15674	24	ABL34477 Human metastasis a
C 22	46	6.6	6668	24	ABL33697 Human immune syste
C 23	45.8	6.5	5195	24	ABL32921 Human immune syste
C 24	45.8	6.5	8592	24	ABL33983 Human immune syste
C 25	45.8	6.5	11650	22	AAS46756 Tumour suppressor
C 26	45.6	6.5	5276	24	ABL32151 Human immune syste
C 27	45.6	6.5	6591	24	ABL33250 Human immune syste
C 28	45.6	6.5	7037	24	ABL33751 Human immune syste
C 29	45.4	6.5	693	22	AAL14804 Human breast cance
C 30	45.4	6.5	5680	24	AAS61047 Human gene regulat
C 31	45.4	6.5	6061	24	ABL32141 Human immune syste
C 32	45.2	6.5	5314	24	ABL32161 Human immune syste
C 33	45.2	6.5	6191	24	ABL33371 Human immune syste
C 34	45.2	6.5	6298	24	ABL32435 Human immune syste
C 35	45.2	6.5	11049	24	ABL32669 Human immune syste
C 36	45.2	6.5	18283	24	AAS61363 Human gene regulat
C 37	45	6.4	6031	22	AAS46621 Tumour suppressor
C 38	45	6.4	6394	24	AAS61278 Tumour suppressor
C 39	45	6.4	6565	22	AAS46465 Tumour suppressor
C 40	45	6.4	7900	22	AAS46760 Human immune syste
C 41	45	6.4	13449	24	ABL33384 Human immune syste
C 42	45	6.4	18154	24	ABL32254 Human immune syste
C 43	44.8	6.4	5474	24	ABL33041 Tumour suppressor
C 44	44.8	6.4	6409	22	AAS46495 Human immune syste
C 45	44.8	6.4	7037	24	ABL33750 Human immune syste

#### ALIGNMENTS

#### RESULT 1

ABL33523/c standard; DNA: 6928 BP.

XX ABL33523;

AC ABL33523;

XX 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1496.

XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytostatic; neutropenic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.

OS Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

```
PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX Claim 1; SEQ ID NO 1496; 32pp + Sequence Listing; German.
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX Sequence 6928 BP; 1852 A; 151 C; 1572 G; 3353 T; 0 other;
Query Match 7.3%; Score 50.8; DB 24; Length 6928;
Best Local Similarity 47.6%; Pred. No. 0.12;
Matches 180; Conservative 0; Mismatches 197; Indels 1; Gaps 1;
QY 304 CTGAGAAATAAGTACTTTGGTCTGAGAAAGACTTAGCGACAACTCCCAATCTCTCA 363
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4902 CTAATACATCAATCTATTTTAAAAAATTAATAAACCTTAATCTAGCCCTTAATTTACTA 4843
QY 364 ATTCCTCAATATATCTTTCCCTTACATGGAATAACTGCATCTTGGATATATCCAAATTA 423
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4842 ATCTATACATCAATCAATTAATACCCAAACAATAAATTTTACTTTTAAAAAACCTTAAT 4783
QY 424 ACACAAATAAATCAACAATAATTTCTTTGAAAGTTCCGGAGAAAACTCAATCTCCACAA 483
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4782 AAACATAAATAACCGTATATAATCTACTATACATAAACAACATAACCAAAAT 4723
QY 484 TTGGATATTGGAGTATTATCGATGGATTTCTTCTTTTAATTAACATGCGCCTATAA 543
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4722 AAAAAATTAAAAAATAAATAATTAATCTTTTAAATTTCTCTTAAATCTATAT-CAACAATAA 4664
QY 544 AAGAAATGGCCACTTGAATATCAAGATCAAAATATTTCAAAATGAAATATGAGCCA 603
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4663 AAAATCACACAAATAAATAATTAATPACAAATTAACCAATAATCTATCAATAC 4604
QY 604 CTAACTCAACAAATCCAGAAATCAGGTGGGAATACTTAACTCAGAACCAACACAAAG 663
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4603 TACTAACTCAAAATCACTTCAACACCATATTATACCTAATAAATAACATAAAACCTCAAA 4544
QY 664 CCTCATCCAGTTGTAGA 681
Db ||| ||| ||| |||
4543 ACCCACACAATTTCTATA 4526
RESULT 2
ABL33561/c
ID ABL33561 standard; DNA; 5814 BP.
XX ABL33561;
XX 26-MAR-2002 (first entry)
XX Human immune system associated gene SEQ ID NO: 1534.
XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
```

```
OS Homo sapiens.
XX WO200200928-A2.
XX 03-JAN-2002.
XX 02-JUL-2001; 2001WO-EP07537.
XX 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX Claim 1; SEQ ID NO 1534; 32pp + Sequence Listing; German.
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX Sequence 5814 BP; 1033 A; 135 C; 1458 G; 3188 T; 0 other;
Query Match 7.1%; Score 49.4; DB 24; Length 5814;
Best Local Similarity 52.6%; Pred. No. 0.24;
Matches 131; Conservative 0; Mismatches 116; Indels 2; Gaps 1;
QY 419 AATTAACACAAATAAATAACCAAAAATTTCTCTTGAAAGTTCCGGAGAAAACTCAATCTC 478
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
681 AAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCGAANAAC 622
QY 479 CACAATTTGGATATTGGAGTATTATCGATGGATTTCTTCTTCTTAAAT--TAAACATGG 536
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
621 AAAAAAATAAATAAATAACCAATACCGGAATATCTCTATCTCTTTAATAACAAAAACAA 562
QY 537 CCTATAAAGAAATGGCCACTTGAATATCAAGATCAAAATATTCAAAATGAAATAT 596
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
561 ACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 502
QY 597 GAAGCCACTAAACTCAACAAATCCAGATCAGGTGGGAATACTTAACCTCAGAACCAAAA 656
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
501 CACACTATATCTCACACTTTAAACCAAAATAAATAAATAAATAAATAAATAAATAAATA 442
QY 657 CACAAGCC 665
Db ||| |||
441 TTCGAAGACC 433
RESULT 3
AAS46771/c
ID AAS46771 standard; DNA; 17280 BP.
XX AAS46771;
XX 18-DEC-2001 (first entry)
XX Tumour suppressor gene derived chemically modified sequence #495.
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
```

OS Homo sapiens.  
PN WO200168912-A2.  
XX  
XX  
XX 20-SEP-2001.  
XX  
XX 15-MAR-2001; 2001WO-EP02955.  
XX  
XX 15-MAR-2000; 2000DE-1013847.  
PR 06-APR-2000; 2000DE-1019058.  
PR 07-APR-2000; 2000DE-1019173.  
PR 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2001-602752/68.  
XX  
XX Fragments of chemically modified genes associated with tumour suppressor  
PT genes and oncogenes, useful in designing primers and probes for  
PT analysing diseases associated with cytosine methylation state e.g.  
PT cancer  
XX  
XX Claim 1; SEQ ID No 495; 27pp; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence of 18  
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
CC bisulphite, of genes associated with tumour suppression and  
CC oncogenes having a sequence taken from 536 (actually 533 since  
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences  
CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a  
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may  
CC form part of a set of probes for detecting the cytosine methylation state  
CC and/or single nucleotide polymorphisms and also to be used in an  
CC array for analysing diseases associated with CpG dinucleotides e.g.  
CC cancers and tumours. The probes can also be used in a method for  
CC ascertaining genetic and/or epigenetic parameters for the diagnosis  
CC and/or therapy of existing diseases or the predisposition to specific  
CC diseases, by analysing cytosine methylations. The parameters may be  
CC compared to another set of genetic and/or epigenetic parameters, the  
CC differences serving as basis for diagnosis and/or prognosis events which  
CC are disadvantageous to patients. The present sequence is one of the  
CC 533 genomic sequences derived from tumour suppressor genes and  
CC oncogenes.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 17280 BP; 5287 A; 151 C; 3639 G; 8203 T; 0 other;  
Query Match 7.0%; Score 48.8; DB 22; Length 17280;  
Best Local Similarity 47.4%; Pred. No. 0.37; Indels 0; Gaps 0;  
Matches 146; Conservative 0; Mismatches 162;  
QY 371 ATATTATCTTTTCCCTTACATGAATACTGCATTCTTGGATATATTCCTCAATTAACACAAA 430  
DB ATAAATATTATTTTATATATATATATTTATATATAAAATATTTATTTATATATATT 8399  
QY 431 TAAATCAACAAATTTCTCTTGAAGTTCGGGAGAAATCAATCTCCCAATTTGGATA 490  
DB TATAACATAAAATATTTATTTATATAATATATTTATATAATAAAATATTTATTTATATA 8339  
QY 491 TTGAGATTATTCATGATGATTTCTTCTTTTAAATTAACATGCGCTTATAAAGAAA 550  
DB TATTTATATAATAAATATTTATTTATATAATAAATATTTATTTATATAAATAATA 8279  
QY 551 TGGCCACTTGAATCAAGAAATCAAAATATTTCAAAATGAATTAATGAAGCCACTTAAACT 610  
DB TTTATTTATATAATAATACTAACTTAAATATAATAAATAAATAAATAAATAAATAA 8219

QY 611 CACAAATCCAGATCAGTGGGATTAACCTTAACCTCAGAACCAAAACACAAAGCCTCATC 670  
DB 8218 CTTCAAAATATATCAAAAAAATAAATTTCCATAATAACTCAAAACAAATAATTTAT 8159  
QY 671 CAGTTGTT 678  
DB 8158 TATTTT 8151  
RESULT 4  
AAS46439/C  
ID AAS46439 standard; DNA; 7040 BP.  
XX  
XX AAS46439;  
XX  
XX 18-DEC-2001 (first entry)  
XX Tumour suppressor gene derived chemically modified sequence #161.  
XX  
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;  
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
KW cytosine methylation; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200168912-A2.  
XX  
XX 20-SEP-2001.  
XX  
XX 15-MAR-2001; 2001WO-EP02955.  
XX  
XX 15-MAR-2000; 2000DE-1013847.  
PR 06-APR-2000; 2000DE-1019058.  
PR 07-APR-2000; 2000DE-1019173.  
PR 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2001-602752/68.  
XX  
XX Fragments of chemically modified genes associated with tumour suppressor  
PT genes and oncogenes, useful in designing primers and probes for  
PT analysing diseases associated with cytosine methylation state e.g.  
PT cancer  
XX  
XX Claim 1; SEQ ID No 161; 27pp; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence of 18  
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
CC bisulphite, of genes associated with tumour suppression and  
CC oncogenes having a sequence taken from 536 (actually 533 since  
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences  
CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a  
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may  
CC form part of a set of probes for detecting the cytosine methylation state  
CC and/or single nucleotide polymorphisms and also to be used in an  
CC array for analysing diseases associated with CpG dinucleotides e.g.  
CC cancers and tumours. The probes can also be used in a method for  
CC ascertaining genetic and/or epigenetic parameters for the diagnosis  
CC and/or therapy of existing diseases or the predisposition to specific  
CC diseases, by analysing cytosine methylations. The parameters may be  
CC compared to another set of genetic and/or epigenetic parameters, the  
CC differences serving as basis for diagnosis and/or prognosis events which  
CC are disadvantageous to patients. The present sequence is one of the  
CC 533 genomic sequences derived from tumour suppressor genes and  
CC oncogenes.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 other;  
SQ Query Match 6.9%; Score 48.4; DB 22; Length 7040;  
Best Local Similarity 52.9%; Pred. No. 0.4;  
Matches 127; Conservative 0; Mismatches 111; Indels 2; Gaps 1;

QY 423 AACCAAAATTAATCAACAAATTTCTGTGAAGTTCGGAGAAAACCAATCTCCACA 482  
DB 1830 AAACATATTTAATAATATTAATTTAAAATAATCTATCAAACTAAAATRACTCG 1771

QY 483 ATGGATATGGAGTATTATCGATGATTTCTTCTTTTAAATTAACATCGCGCTATA 542  
DB 1770 TTTTCATATAACATCAACAATTAATATTTTCTTCTTAAATTTTACGATTA 1711

QY 543 AAGAAAATGGCCACTTGGAATATCAAGAAATCAAAATTTCAAAATGAATGAAGCC 602  
DB 1710 AAATATAAATCCTACTATATTAATAATACACATATATAAAACGTAATAAATAAAT 1651

QY 603 ACTAACTCAACAATCCAGAAATCAGGTGGGAATACCTTAACCTCAGAACCAACAAA 562  
DB 1650 A--ACACTAATCAATAACAATATATACCTATATATATCCCTCCCAAAAACAAAACAAA 1593

RESULT 5  
ABL33368/c  
ID ABL33368 standard; DNA; 9817 BP.  
XX AC ABL33368;  
XX DT 26-MAR-2002 (first entry)  
XX DE Human immune system associated gene SEQ ID NO: 1341.  
XX DE Human immune system associated gene SEQ ID NO: 1341.  
KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytotatic; neutropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX OS Homo sapiens.  
XX XX WO200200928-A2.  
XX PN 03-JAN-2002.  
XX PD 02-JUL-2001; 2001WO-EP07537.  
XX PF 30-JUN-2000; 2000DE-1032529.  
XX PR 01-SEP-2000; 2000DE-1043826.  
XX XX (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX XX WPI; 2002-130909/17.  
XX DR Nucleic acid comprising fragment of chemically modified gene, useful  
XX PT for diagnosis and treatment of diseases associated with abnormal  
XX PT cytosine methylation -  
XX PS Claim 1; SEQ ID NO 1341; 32pp + Sequence Listing; German.  
XX CC The present invention provides a number of human immune system associated  
XX CC genes which are modified by the methylation of cytosines. The sequences  
XX CC can be used in the diagnosis and treatment of immune system disorders,  
XX CC including eye diseases such as retinopathy, neovascular glaucoma and  
XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
XX CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

CC diseases. The present sequence is a gene of the invention.  
XX SQ Sequence 9817 BP; 2814 A; 174 C; 2033 G; 4794 T; 2 other;  
Query Match 6.9%; Score 48.2; DB 24; Length 9817;  
Best Local Similarity 48.4%; Pred. No. 0.46;  
Matches 134; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 354 CAATCTCTTAATCTCGAATATATCTTTTCCCTACATGGAATAACTCTCTTGATAT 413  
DB 1839 CAACCTTTAAATTCATATTCATTAATAAATAATAATATACATATCTACTTACT 1780

QY 414 ATTCATTAATTAACACAAATAAATAAATAAATTTCTTGAAGTTCGGAGAAACTCA 473  
DB 1779 ACTAAAATTTAAACACACATTAATCTCTTAAAAAATTTCTACTCAAAAAATAA 1720

QY 474 ATCTCCACAAATGGATATTTGGAGTATATTCGATGGATTTTCTTCTTAAATTAACAT 533  
DB 1719 ATTTCTAAATTAATCTCAAAACCTTAATACAAATAAATTAACATAATTAATCACTCAAAA 1660

QY 534 GCGCCTATATAAAGAAAATGGCCACTTGAATAATCAAGAAATCAAAATATTTCAAAATGAAT 593  
DB 1659 AAACCTTAATAACATTAATCTACACGAAATAAATAATAATATATATACATTAATAAATAA 1600

QY 594 AATGAAGCCACTAAACTCAACAAATCCAGATCAGGT 630  
DB 1599 AATAATTCATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1563

RESULT 6  
ABL33414  
ID ABL33414 standard; DNA; 17389 BP.  
XX AC ABL33414;  
XX DT 26-MAR-2002 (first entry)  
XX DE Human immune system associated gene SEQ ID NO: 1387.  
XX DE Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytotatic; neutropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX OS Homo sapiens.  
XX XX WO200200928-A2.  
XX PN 03-JAN-2002.  
XX PD 02-JUL-2001; 2001WO-EP07537.  
XX PF 30-JUN-2000; 2000DE-1032529.  
XX PR 01-SEP-2000; 2000DE-1043826.  
XX XX (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX XX WPI; 2002-130909/17.  
XX DR Nucleic acid comprising fragment of chemically modified gene, useful  
XX PT for diagnosis and treatment of diseases associated with abnormal  
XX PT cytosine methylation -  
XX PS Claim 1; SEQ ID NO 1387; 32pp + Sequence Listing; German.  
XX CC The present invention provides a number of human immune system associated  
XX CC genes which are modified by the methylation of cytosines. The sequences



CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.  
XX  
SQ Sequence 17389 BP; 4806 A; 201 C; 3688 G; 8694 T; 0 other;

Query Match 6.9%; Score 48.2; DB 24; Length 17389;  
Best Local Similarity 47.0%; Pred. No. 0.49;  
Matches 149; Conservative 0; Mismatches 168; Indels 0; Gaps 0;  
QY 14 GATTTTTCCTTTTACCTATTTCAATTTAGTTTCTTTGATTCACAAAGTC 73  
DB 779 GATTTTTCCTTTTACCTATTTCAATTTAGTTTCTTTGATTCACAAAGTC 838  
QY 74 ATTATTTGTTTAAATCCAGATCGGCAATTTTGAATAAATTTCTACTAAATTCGA 133  
DB 839 TTTAGAGGTATGTTAAATTTCTTAAATTTATTAATGAAGTTGGGTTATAGGTA 898  
QY 134 AATTAATTTGACCTTTGGCTTGATAATGCTGCTCGGAGGTGATATTTTATCCT 193  
DB 899 ATTTGAATTTTATTTTCTTTGTTTGGTTAGATCTTTAGTAATATTAATATTTAT 958  
QY 194 TGATAGAAACAGCTCTGCAAGCTGTAAGTTATTCATCGCTGTGTTTCTATACAA 253  
DB 959 TTAATTTGGAATAATGTTTAAATTTATTCGATTTATTTTATTTATGATATAAT 1018  
QY 254 AACATCTTTTGAAGAATGAACACATTCCTTAAACCCCTTAAATCGCGCTGAGATA 313  
DB 1019 ATTATGTTGGATGATATTTATTTGTTATGTTAGATTTTGTGATGATGAATAA 1078  
QY 314 AATCAGCTTTTGGTCTG 330  
DB 1079 AATGATGCTGTTTG 1095

RESULT 7  
ABL32335/C  
ID ABL32335 standard; DNA; 10528 BP.

XX ABL32335;  
XX  
XX  
XX 26-MAR-2002 (first entry)  
DE Human immune system associated gene SEQ ID NO: 308.

XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; anti-anaemic; cytosine; nontropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.

XX Homo sapiens.  
OS  
XX  
XX WO200200928-A2.  
PN  
XX  
XX 03-JAN-2002.  
PD  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
PF  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
PR  
XX  
XX 01-SEP-2000; 2000DE-1043826.  
PR  
XX  
XX (EPIC-) EPIGENOMICS AG.  
PA  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
PI  
XX  
XX WPI; 2002-130909/17.

XX  
PT  
PT  
XX  
PS  
XX  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
SQ

Nucleic acid comprising fragment of chemically modified gene, useful  
for diagnosis and treatment of diseases associated with abnormal  
cytosine methylation -

Claim 1; SEQ ID NO 308; 32pp + Sequence Listing; German.

The present invention provides a number of human immune system associated  
genes which are modified by the methylation of cytosines. The sequences  
can be used in the diagnosis and treatment of immune system disorders,  
including eye diseases such as retinopathy, neovascular glaucoma and  
macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
diseases. The present sequence is a gene of the invention.

Sequence 10528 BP; 2873 A; 86 C; 2164 G; 5405 T; 0 other;

Query Match 6.9%; Score 48; DB 24; Length 10528;  
Best Local Similarity 50.0%; Pred. No. 0.51;  
Matches 146; Conservative 0; Mismatches 145; Indels 1; Gaps 1;  
QY 326 TTCTGAGAAAGACTTAGCGACAACCTCTCCAATCTTCTTAATTTCTCAATATTTTCTCT 385  
DB 4679 TTCTTAAATCAATCCTTAATATTAATTAACCTTATAATTTCTTATTAATAAATTTTA 4620  
QY 386 TACATGGAATTAATGCAATCTTGGATATATTCCAATTTAACACAAATTAACACAAAT 445  
DB 4619 TATACCTCATTAACAACTTTATTAATAGCTCTCTTTAAC-CAATACTATCATATAAAC 4561  
QY 446 TTCTCTGGAAGTTCGGAGAAAACCTCAATCTCCACAATTTGGATATTTGGAGTATTATCGA 505  
DB 4560 TAAAAATTAACAACTAATCAATCCTTAAATCTTAAAAAACATCTATATATAACAAATAA 4501  
QY 506 TGGATTTTCTTCTTCTTAAATTAACATCGCCCTTATAAAGAAATGCCACTTTGAATAA 565  
DB 4500 CATATAAATACTTCTTAATACATACATATATATATATAAATAATATATACAAAAA 4441  
QY 566 TCAAGAAATCAAAATTTTCAAAATGAAATATGAAGCCCACTAAACCTCAACAAA 617  
DB 4440 ATCAAAATTAACCTTCAAAAAAATAAATTTTAAACAAAAATATTTAAAAA 4389

RESULT 8

AAT31530/C  
ID AAT31530 standard; cDNA; 605 BP.

XX AAT31530;  
XX  
XX 15-SEP-1996 (first entry)  
DE Human 3' apolipoprotein B SAR element clone Rh32.  
DE  
DE  
DE  
KW Erythropoietin; EPO; anaemia; gene therapy; vector;  
KW scaffold attachment region; SAR element; apolipoprotein B;  
KW transgenic animal; ss.

XX Homo sapiens.

OS  
XX  
XX WO9619573-A1.

PN  
XX  
XX 27-JUN-1996.

PD  
XX  
XX 18-DEC-1995; 95WO-CA00696.

PF  
XX  
XX 19-DEC-1994; 94US-0358918.

PR  
XX  
XX (CANG-) CANGENE CORP.

PA  
XX  
XX Delcuve G;

PI  
XX  
XX WPI; 1996-309587/31.

DR  
XX



XX PF 06-APR-2001; 2001WO-EP03968.  
XX PR 06-APR-2000; 2000DE-1019058.  
XX PR 07-APR-2000; 2000DE-1019173.  
XX PR 30-JUN-2000; 2000DE-1032529.  
XX PR 01-SEP-2000; 2000DE-1043826.  
XX PA (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX DR WPI; 2002-017470/02.  
XX PT New nucleic acid sequences from chemically modified genes associated  
XX PT with gene regulation, useful for analysing cytosine methylations for  
XX PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency  
XX PT disease -  
XX PS Claim 1; SEQ ID No 204; 26pp; English.  
XX CC The invention relates to 224 nucleic acid sequences comprising at least  
XX CC 18 bases of a chemically pretreated gene associated with gene regulation  
XX CC selected from 43 known genes (or complementary sequences). The  
XX CC chemical pretreatment converts cytosine bases unmethylated at the  
XX CC 5-position to uracil or another base with hybridisation behaviour  
XX CC dissimilar to cytosine, to enable analysis of cytosine methylations.  
XX CC The DNA sequences, oligomers (or sets/arrays) and method are  
XX CC useful in the diagnosis of diseases (or predisposition to diseases)  
XX CC associated with gene regulation and in therapy of such diseases, by  
XX CC enabling analysis of the cytosine methylation patterns of such genes,  
XX CC kits are provided. They are especially useful in diagnosis  
XX CC and therapy of e.g. severe combined immunodeficiency disease, cardiac  
XX CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,  
XX CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,  
XX CC pre-eclampsia, graft versus-host disease. The present sequence is a  
XX CC sequence included in the sequence data for this specification and is  
XX CC associated with the human gene regulation-associated genes.  
XX CC Note: The sequence data for this patent did not form part  
XX CC of the printed specification, but was obtained in electronic  
XX CC format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SO Sequence 5501 BP; 1484 A; 107 C; 1262 G; 2648 T; 0 other;  
Query Match 6.7%; Score 47.2; DB 24; Length 5501;  
Best Local Similarity 46.6%; Pred. No. 0.68;  
Matches 151; Conservative 0; Mismatches 173; Indels 0; Gaps 0;  
QY 334 AAGACTTAGCGACAACTCTCGAATCTTCTTAATCTCGAATATATATCTTTCTTACATGGA 393  
Db 2933 AAAATATAAAAAAACAATTAATAAATAAATTAATAAATAAATAAATAAATAAATAA 2874  
QY 394 ATAATCGATCTTGGATATATTCCTAATTAACACAAATAAATAAATAAATAAATAAATAA 453  
Db 2873 TTAATACTATTAACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2814  
QY 454 AAAGTTCGGAGAAACTCAATCTCCCAATTTGGATATTTGGAGTATTATCGATGATTTT 513  
Db 2813 AAAATCTTTAAACATTTTATCTTTAAATACCTATTATATAAATAAATAAATAAATAA 2754  
QY 514 CTTCCTCTTTAAATTAACATGCGCTATTAAGAAATAATGGCCACTTGAATATCAAGANT 573  
Db 2753 ACAACCTTAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2694  
QY 574 CAAATATTTCAAAATGAATTAATGAAGCCCTAACTCAACAAATCCAGATCAGGTGGG 633  
Db 2693 ANAACTTTCAAACTAAACATACCATTAAACAAATAAATAAATAAATAAATAAATAA 2634  
QY 634 AATACTTAACCTCAGAACCAAAAC 657  
Db 2633 AATTAAAAAATAAATAAATAA 2610

## RESULT 11

ABL34428/c  
ID ABL34428 standard; DNA; 9110 BP.  
XX AC ABL34428;  
XX DT 26-MAR-2002 (first entry)  
XX XX Human immune system associated gene SEQ ID NO: 2401.  
XX DE Human; immune system disease; cytosine methylation; antiasthmatic;  
XX KW antiarteriosclerotic; antianemic; cytostatic; nootropic;  
XX KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
XX KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
XX KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
XX KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
XX KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
XX KW gene; ds.  
XX OS Homo sapiens.  
XX PN WO200200928-A2.  
XX PD 03-JAN-2002.  
XX PF 02-JUL-2001; 2001WO-EP07537.  
XX PR 30-JUN-2000; 2000DE-1032529.  
XX PR 01-SEP-2000; 2000DE-1043826.  
XX XX (EPIG-) EPIGENOMICS AG.  
XX XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-130909/17.  
XX PT Nucleic acid comprising fragment of chemically modified gene, useful  
XX PT for diagnosis and treatment of diseases associated with abnormal  
XX PT cytosine methylation -  
XX PS Claim 1; SEQ ID NO 2401; 32pp + Sequence Listing; German.  
XX CC The present invention provides a number of human immune system associated  
XX CC genes which are modified by the methylation of cytosines. The sequences  
XX CC can be used in the diagnosis and treatment of immune system disorders,  
XX CC including eye diseases such as retinopathy, neovascular glaucoma and  
XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
XX CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
XX CC diseases. The present sequence is a gene of the invention.

XX SO Sequence 9110 BP; 2699 A; 146 C; 2072 G; 4193 T; 0 other;

Query Match 6.7%; Score 47.2; DB 24; Length 9110;  
Best Local Similarity 46.6%; Pred. No. 0.73;  
Matches 151; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 334 AAGACTTAGCGACAACTCTCCAAATCTTCTTAATCTCGAATATATATCTTTCTTACATGGA 393  
Db 6542 AAAATATAAAAAAACAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6483  
QY 394 ATAATCGATCTTGGATATATTCCTAATTAACACAAATAAATAAATAAATAAATAAATAA 453  
Db 6482 TTAATACTATTAACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6423  
QY 454 AAAGTTCGGAGAAACTCAATCTCCCAATTTGGATATTTGGAGTATTATCGATGATTTT 513  
Db 6422 AAAATCTTTAAACATTTTATCTTTAAATACCTATTATATAAATAAATAAATAAATAA 6363  
QY 514 CTTCCTCTTTAAATTAACATGCGCTATTAAGAAATAATGGCCACTTGAATATCAAGANT 573  
Db 6362 ACAACCTTAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6303





Qy	301	CGCGTGAAGATAAAATGACTTTTGGTCTTGAGAAGACTTTAGCGACAACTCTCCAATCTTT	360
Db	27535	ATACTAAAAATAAAAACAAAAATACAATTTCTACTAAATATCTACTACTTTTCACACCAT	27476
Qy	361	CTAATTCGGAATATATCTTTTCCATTACATGGAATAACTGCAATTCCTTGGATATATTTCCA -	419
Db	27475	CGTAAATTAATAAATCCATAATCAAAACCATCAACAATCAGAAACTTATATATACTTTCTAT	27416
Qy	420	--ATTAAACACAAATATAAATCAACAAAAATTTCTCTTGAAAGTTCCGGAGAAAACTCAATCT	477
Db	27415	TTATAATTAACCAATTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCT	27356
Qy	478	CCACAATTTGGATATTGGAGTATTTATCGATGGAATTTCTCTCTTTAACTAAACATCGCG	537
Db	27355	ATACCAATTAATTTACAAAAAATCTATTTAAAAAACCACATCCCTCAACATAAATAAAAAA	27296
Qy	538	CTATATAAAGAAAAATGGCCACTTTGGAATTAATCAAGAAATCAAAATATTTCAAAATGAATG	597
Db	27295	AATAACAAAAAACAACACCATTTATATAATTTTAAACAATAAAAAAATAAAAAACTAAAA	27236
Qy	598	AAGCCACTTAACTCAACAAATCCAGAATTCAGTGGGAATTAACCTTAACCTCAGAACCAAAAC	657
Db	27235	ACTAAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	27176
Qy	658	ACAAAGCCTCA	668
Db	27175	AAAAAAACTAA	27165

Search completed: November 1, 2002, 23:23:04  
Job time : 282.968 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:17:03 ; Search time 34.9206 Seconds  
(without alignments)  
4923.835 Million cell updates/sec

Title: US-09-881-556-3  
Perfect score: 700  
Sequence: 1 gaattcgccagcagattttt.....accgcattctacagaaaagc 700

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents-NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43.6	6.2	665	2	US-08-883-795A-36
C 2	40.6	5.8	3515	2	US-08-391-743A-1
C 3	40	5.7	4163	4	US-09-004-838-70
C 4	40	5.7	4208	4	US-09-004-838-1
C 5	39.6	5.7	3706	4	US-08-913-159-9
C 6	39.4	5.6	1104	1	US-08-202-186-17
C 7	39.4	5.6	1105	1	US-08-202-186-18
C 8	39.2	5.6	176373	3	US-09-128-155-17
C 9	38	5.4	6138	4	US-09-067-800-4
C 10	38	5.4	6138	4	US-09-349-677-4
C 11	37.8	5.4	1103	1	US-08-202-186-16
C 12	37.8	5.4	1110	1	US-08-202-186-14
C 13	37.8	5.4	2030	2	US-08-705-937-7
C 14	37.4	5.3	10660	2	US-08-267-803B-8
C 15	37.4	5.3	10660	4	US-09-041-886-16
C 16	37	5.3	9636	1	US-08-323-170B-1
C 17	37	5.3	9636	4	US-08-954-441-1
C 18	36.8	5.3	1431	4	US-09-316-083-2
C 19	36.8	5.3	1539	4	US-07-956-483-21
C 20	36.8	5.3	1539	4	US-08-472-240A-13
C 21	36.8	5.3	2571	1	US-08-254-358-3
C 22	36.8	5.3	2571	1	US-08-475-391-3
C 23	36.8	5.3	2571	2	US-08-709-609-3
C 24	36.8	5.3	2571	5	PCT-US95-07178-3
C 25	36.8	5.3	3807	2	US-08-417-210A-78
C 26	36.8	5.3	4529	2	US-08-449-645A-16
C 27	36.8	5.3	4529	2	US-08-702-367A-16

C	28	36.8	5.3	4529	5	PCT-US95-04681-16	Sequence 16, Appl
	29	36.8	5.3	6474	3	US-08-651-472-66	Sequence 66, Appl
	30	36.8	5.3	6474	4	US-08-358-928-66	Sequence 66, Appl
C	31	36.8	5.3	6926	3	US-08-651-472-69	Sequence 69, Appl
C	32	36.8	5.3	6926	4	US-08-358-928-69	Sequence 69, Appl
C	33	36.8	5.3	9739	1	US-08-022-835-1	Sequence 1, Appl
C	34	36.8	5.3	9739	1	US-08-388-809-1	Sequence 1, Appl
C	35	36.8	5.3	9739	2	US-08-647-714-1	Sequence 1, Appl
C	36	36.8	5.3	9746	1	US-08-022-835-3	Sequence 3, Appl
C	37	36.8	5.3	9746	1	US-08-388-809-3	Sequence 3, Appl
C	38	36.8	5.3	9746	2	US-08-647-714-3	Sequence 3, Appl
C	39	36.6	5.2	615	4	US-08-998-416-186	Sequence 186, App
C	40	36.6	5.2	4700	2	US-08-828-692-16	Sequence 16, Appl
C	41	36.4	5.2	665	2	US-08-883-795A-36	Sequence 36, Appl
C	42	36.4	5.2	3138	1	US-07-867-106-4	Sequence 4, Appl
C	43	36.2	5.2	396	1	US-08-470-179-48	Sequence 48, Appl
C	44	36.2	5.2	982	3	US-08-973-068-28	Sequence 28, Appl
C	45	36.2	5.2	1109	1	US-08-202-186-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-08-883-795A-36/c  
; Sequence 36, Application US/08883795A  
; Patent No. 5985607  
; GENERAL INFORMATION:  
; APPLICANT: Delcove, Genevieve  
; APPLICANT: Awang, Gregor  
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression  
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERESKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/883,795A  
; FILING DATE: 27-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gravelle, Micheline  
; REGISTRATION NUMBER: 40,261  
; REFERENCE/DOCKET NUMBER: 7841-062  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 665 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; IMMEDIATE SOURCE:  
; CLONE: Rh 32  
; US-08-883-795A-36

Query Match 6.2%; Score 43.6; DB 2; Length 665;  
Best Local Similarity 49.6%; Pred. No. 0.091;  
Matches 112; Conservative 0; Mismatches 114; Indels 0; Gaps 0;



QY 383 CCTTACATGAATACTGCATCTTGGATATATTCCTCAATTAACACAAATAAATCAACAA 442  
Db 397 CATTTTAATTAATAATATGTATTAATAACATTTTAAATTAATAAATATGTAATTATAAA 338  
QY 443 AATTTCTCTTGAAGTTCGGGAGAAAACCTCAATCTCCACAATGGGATATTGGAGTATTAT 502  
Db 337 CATTTTAATTAATAATATGTATTAATAACATTTTAAATTAATAAATATTTAATTATAAA 278  
QY 503 CGATGGATTTCCTCTCTCTTAATTAACATCGGCTATATAAAGAAAATGGCCACTTGAA 562  
Db 277 CATTTTAATTAATAAATAATTAATAAATAATTTTAAATTAATAAATAATTTAATTATAAA 218  
QY 563 TAATCAAGATCAAAATATTCCTCAATGAATGAATGAATGAATGAATGAATGAATGAAT 608  
Db 217 TATTTTAATTAATAAATAATTTAATAAATAATTTAATAAATAATTTAATAAATAA 172

RESULT 2  
US-08-391-743A-1  
; Sequence 1, Application US/08391743A  
; Patent No. 5843705  
; GENERAL INFORMATION:  
; APPLICANT: DiTullio, Paul A.; Meade, Harry; Cole, Edward S.  
; TITLE OF INVENTION: TRANSGENETICALLY PRODUCED ANTI-THROMBIN III  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/391.743A  
; FILING DATE: 21-FEB-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Paul Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: TCI-045  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3515 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: YES  
; FRAGMENT TYPE: N-terminal  
US-08-391-743A-1

Query Match 5.8%; Score 40.6; DB 2; Length 3515;  
Best Local Similarity 48.1%; Pred. No. 0.61;  
Matches 115; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 307 AGAATAAATGACTTTTGGTCTGAGAAGACTTAGGACAACTCTCAATCTTCTTAAT 366  
Db 1404 ATAAACAACTGTATTATCTTTTGGTTACATTAAATGGCAACCACCTCCAGTACTCTTG 1463  
QY 367 CTGATATATATCTTTCTCTACATGGAATAACTGCATCTCTGGATATATTCGAATTAACA 426

Db 1464 ACAATATGAATTTTCTTTAAAGCTAAACCTGATTTTATTTTATTTTTCACAAAGGAAT 1523  
QY 427 CAATAAATAATCAACAAATTTCTCTTGAAGTTCGGGAGAAAACCTCAATCTCCACAATG 486  
Db 1524 CTATTACACACATCAATRAAGGTAAACCTTCATATTAAACGTATCATTTTAAAAATTT 1583  
QY 487 GATATTGGAGTATTATCGATGGATTTTCTCTCTTTAATTAACATGCGCCTATATAAA 545  
Db 1584 CATGTTTGAATTTTATAAACAGCATTTCTTTATGTGTCATTTTTTTTTTACCAGAAA 1642

RESULT 3  
US-09-004-838-70  
; Sequence 70, Application US/09004838  
; Patent No. 6350933  
; GENERAL INFORMATION:  
; APPLICANT: Michelmore, Richard W.  
; APPLICANT: Shen, Kathy  
; APPLICANT: Meyers, Blake  
; TITLE OF INVENTION: Procedures and Materials for  
; TITLE OF INVENTION: Conferring Pest Resistance in Plants  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/004.838  
; FILING DATE: 09-JAN-1998  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/781,734  
; FILING DATE: 10-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Einhorn, Gregory P.  
; REGISTRATION NUMBER: 38,440  
; REFERENCE/DOCKET NUMBER: 023070-078810US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 70:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4163 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; FEATURE:  
; NAME/KEY:  
; LOCATION: 1..4163  
; OTHER INFORMATION: /note= "RLG1-EL69"  
US-09-004-838-70

Query Match 5.7%; Score 40; DB 4; Length 4163;  
Best Local Similarity 52.9%; Pred. No. 0.87;  
Matches 108; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 393 AATAACTGCATCTTGGATATATTCCTCAATTAACAAATAAATAAATAAATTTCTCTT 452  
Db 177 AATAAATGCATATTTATTTTCTTTAAATAAACGCATATATATAGATTA-AAATCAT 235  
QY 453 GAAAGTTCGGAGAAAACCTCAATCTCCCAATTCGATATTCGAGTATATATCGATGGATT 512  
Db 236 ATAATACATAGGTAAACTCATATATATATATGTTTCATCCCAAGTTTATTTATATGCT 295



Db 619 ATAGGGGAGGACTTTTGCTTTCACAAAAAATATCTTAAAAATTTTCAATTCACAT 560  
QY 459 TCGGGAACAACTCAATCTCCACAAATGGATATTTGAGTATTATCGATGGATTTTCTTCT 518  
Db 559 TTTGAAAAGACGAATTAATCTTTTATTAATAATGGATCTTAAAGAGAGCTTGATTA 500  
QY 519 TCTTTAATTAACATCGGCTATATAAGAAATGGCCACTTGAATATCAAGAAATCAA 576  
Db 499 TATTTAATGAAAAATACCAATTAATATAATAATATGAAGAGTATATGAATAACGAA 442

## RESULT 6

US-08-202-186-17/c

; Sequence 17, Application US/08202186

; Patent No. 5756708

; GENERAL INFORMATION:

; APPLICANT: KARAN, Mirko

; APPLICANT: BURNS, Thomas M.

; APPLICANT: DALE, James L.

; APPLICANT: HARDING, Robert M.

; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley &amp; Lardner

; STREET: 3000 K Street, N.W.

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/202,186

; FILING DATE: 24-FEB-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: JEFFERY, Donald D.

; REGISTRATION NUMBER: 19,980

; REFERENCE/DOCKET NUMBER: 71611/102 FIKE

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202 672 5300

; TELEFAX: 202 672 5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1104 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: circular

US-08-202-186-17

Query Match 5.6%; Score 39.4; DB 1; Length 1104;

Best Local Similarity 53.6%; Pred. No. 0.92;

Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 379 TTTTCTTACATGGAATACTGCATCTTGGATATATCCAAATTAACACAAATAAAATCA 438  
Db 854 TATTTCCCGCTTGGATATATCCATTTTAAATTTCTTAATAAACCATAGTTTAAATAT 795  
QY 439 ACAAAATTTCTTGAAGTTCCGGAGAAACTCAATCTCCACAATTTGGATATTTGGAGTA 498  
Db 794 TCTCTTTGCATCTGGGAATCAAAATATACTATATCTCTCATAATATACAATCTACAT 735  
QY 499 TTATCGATGATTTTCTTCTTTTAAATTAAC 531  
Db 734 ATATCCAATGATTTTCTCTCTCGCGGAATAAAAC 702

## RESULT 7

US-08-202-186-18/c

; Sequence 18, Application US/08202186  
; Patent No. 5756708  
; GENERAL INFORMATION:  
; APPLICANT: KARAN, Mirko  
; APPLICANT: BURNS, Thomas M.  
; APPLICANT: DALE, James L.  
; APPLICANT: HARDING, Robert M.  
; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/202,186  
; FILING DATE: 24-FEB-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JEFFERY, Donald D.  
; REGISTRATION NUMBER: 19,980  
; REFERENCE/DOCKET NUMBER: 71611/102 FIKE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 672 5300  
; TELEFAX: 202 672 5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1105 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: circular  
US-08-202-186-18

Query Match 5.6%; Score 39.4; DB 1; Length 1105;

Best Local Similarity 53.6%; Pred. No. 0.92;

Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 379 TTTTCTTACATGGAATACTGCATCTTGGATATATCCAAATTAACACAAATAAAATCA 438  
Db 855 TATTTCCCGCTTGGATATATCCATTTCTTAATAAACCATAGTTTAAATAT 796  
QY 439 ACAAAATTTCTTGAAGTTCCGGAGAAACTCAATCTCCACAATTTGGATATTTGGAGTA 498  
Db 795 TCTCTTTGCATCTGGGAATCAAAATATACTATTTCTCATAATATACAATCTACAT 736  
QY 499 TTATCGATGATTTTCTTCTTTTAAATTAAC 531  
Db 735 ATATCCAATGATTTTCTCTCTCGCGGAATAAAAC 703

## RESULT 8

US-09-128-155-17/c

; Sequence 17, Application US/09128155

; Patent No. 6117654

; GENERAL INFORMATION:

; APPLICANT: Pan, Yang

; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY

; FILE REFERENCE: 09404/052001

; CURRENT APPLICATION NUMBER: US/09/128,155

; CURRENT FILING DATE: 1998-08-03

; EARLIER APPLICATION NUMBER: US 60/091,650

; EARLIER FILING DATE: 1998-07-02

; EARLIER APPLICATION NUMBER: US 60/054,646

; EARLIER FILING DATE: 1997-08-04

; NUMBER OF SEQ ID NOS: 18



Db 3679 GGTGAATGACATGAGCTGTTTAACTGGAACGTATAGAAAACGGGAATGACAAGT 3620  
QY 538 CTATAAGAAATGGCCACTTGAATAATCAAGAAATCAAAATATTC 583  
Db 3619 CCCCAACTGAGTGGAGAAATAACAATCAAAAATCAAAATTTTC 3574

## RESULT 11

US-08-202-186-16/c  
; Sequence 16, Application US/08202186  
; Patent No. 5756708  
; GENERAL INFORMATION:  
; APPLICANT: KARAN, Mirko  
; APPLICANT: BURNS, Thomas M.  
; APPLICANT: DALE, James L.  
; APPLICANT: HARDING, Robert M.  
; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/202,186  
; FILING DATE: 24-FEB-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JEFFERY, Donald D.  
; REGISTRATION NUMBER: 19,980  
; REFERENCE/DOCKET NUMBER: 71611/102 FIRE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 672 5300  
; TELEFAX: 202 672 5399  
; TELEX: 904136

INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1103 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: circular  
US-08-202-186-16

Query Match 5.4%; Score 37.8; DB 1; Length 1103;  
Best Local Similarity 52.9%; Pred. No. 2.1;  
Matches 81; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 379 TTTTCTTACATGGAATGCAATCTTGGATATATTCCTCAATTAACACAAATAAATCA 438  
Db 853 TATTTCCCGCTTGAATATTCATTTTAAATCTCTTAAGCCATAGTTTAATAT 794  
QY 439 ACAAAATTTCTCTGAAAGTCCGGAGAAATCAATCTCCACAATTTGGATATTTGGAGTA 498  
Db 793 TCCTCTTTGCATCTGGATATCAATATCAATATCTCTCATATTAATATACATCTACAT 734  
QY 499 TTATCGATGGATTTCTTCTTTTAAATTAAC 531  
Db 733 ATATCCAATGATTTTCTCTCTCGGCAATAAAC 701

## RESULT 12

US-08-202-186-14/c  
; Sequence 14, Application US/08202186  
; Patent No. 5756708  
; GENERAL INFORMATION:  
; APPLICANT: KARAN, Mirko  
; APPLICANT: BURNS, Thomas M.

; APPLICANT: DALE, James L.  
; APPLICANT: HARDING, Robert M.  
; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/202,186  
; FILING DATE: 24-FEB-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JEFFERY, Donald D.  
; REGISTRATION NUMBER: 19,980  
; REFERENCE/DOCKET NUMBER: 71611/102 FIRE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 672 5300  
; TELEFAX: 202 672 5399  
; TELEX: 904136

INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1110 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: circular  
US-08-202-186-14

Query Match 5.4%; Score 37.8; DB 1; Length 1110;  
Best Local Similarity 52.9%; Pred. No. 2.1;  
Matches 81; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 379 TTTTCTTACATGGAATGCAATCTTGGATATATTCCTCAATTAACACAAATAAATCA 438  
Db 852 TATTTCCCGCTTGAATATTCATTTTAAATCTCTTAATAACCAATTAATTAATAA 793  
QY 439 ACAAAATTTCTCTGAAAGTCCGGAGAAATCAATCTCCACAATTTGGATATTTGGAGTA 498  
Db 792 TCCTCTTTGCATCTAGGATATCAATATCAATATCTCTCATATTAATATACATCTACAT 733  
QY 499 TTATCGATGGATTTCTTCTTTTAAATTAAC 531  
Db 732 ATATCCAATGATTTTCTCTCTCGGCAATAAAC 700

## RESULT 13

US-08-705-937-7  
; Sequence 7, Application US/08705937  
; Patent No. 5981841  
; GENERAL INFORMATION:  
; APPLICANT: Santino, Colleen G.  
; APPLICANT: Conner, Timothy W.  
; TITLE OF INVENTION: EARLY SEED 5' REGULATORY SEQUENCE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ANTORWAY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 10660 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 936..3384
US-09-041-886-16

Query Match      5.3%; Score 37.4; DB 4; Length 10660;
Best Local Similarity 60.2%; Pred. No. 4.1;
Matches 62; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 515 TTCTTCTTTTAAATGAACATGCGCCTATATAAGAAATGCGCCACTTGAATAATCAAGAATC 574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10557 TTATGTTTCAGTATCTCGTACTAATAAATAACAGTGCCCAATGCAAAAAAAAAA 10616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 575 AAAATATTCAAAATGAATAATGAAGCCCACTAAACTCAACAAA 617
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10617 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 10659
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: November 1, 2002, 23:26:13
Job time : 264.421 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:27:34 ; Search time 1235.45 Seconds  
(without alignments)  
7647.311 Million cell updates/sec

Title: US-09-881-556-3

Perfect score: 700

Sequence: 1 gaattggcagcagattttt.....accgatctctacagaaagc 700

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: gb\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	64.2	9.2	1101	12	CNS0039G	AL063921 Drosophill
2	61	8.7	1092	12	CNS020K7	AL175696 Tetraodon
3	55.2	7.9	1190	12	CNS020N7	AL206908 Tetraodon
4	54.8	7.8	836	12	CNS01100	AL099642 Drosophill
5	54.8	7.8	928	12	CNS00DKY	AL071865 Drosophill
6	54.8	7.8	1101	12	CNS0100X	AL098379 Drosophill
7	54.6	7.8	1101	12	CNS000B8	AL063632 Drosophill
8	54.6	7.8	1101	12	CNS00EVL	AL069706 Drosophill
9	53.8	7.7	1101	12	CNS002FG	AL062437 Drosophill
10	53.8	7.7	1101	12	CNS0145U	AL103740 Drosophill
11	53.2	7.6	1101	12	CNS0039G	AL063921 Drosophill
12	52.4	7.5	1101	12	CNS00EVL	AL069706 Drosophill
13	52.4	7.5	1101	12	CNS017KX	AL108171 Drosophill
14	52.2	7.5	1043	12	CNS0145P	AL103735 Drosophill
15	52.2	7.5	1200	12	CNS016CO	AL106578 Drosophill
16	51.8	7.4	1101	12	CNS003BD	AL064091 Drosophill
17	51.6	7.4	867	10	BE782580	601465880

C 18	51.6	7.4	1101	12	CNS00EPYG	AL071206 Drosophill
C 19	51.4	7.3	987	12	CNS00K3Q	AL077186 Drosophill
C 20	51.2	7.3	759	12	CNS007BZ	AL066974 Drosophill
C 21	51.2	7.3	1101	12	CNS00FVE	AL071298 Drosophill
C 22	51.2	7.3	1200	12	CNS007EV	AL067078 Drosophill
C 23	51	7.3	759	12	CNS006XV	AL411257 T7 end of
C 24	51	7.3	1101	12	CNS003DX	AL064587 Drosophill
C 25	51	7.3	1201	12	CNS0167M	AL106396 Drosophill
C 26	50.6	7.2	1091	12	CNS014AC	AL103902 Drosophill
C 27	50.6	7.2	1101	12	CNS00FMC	AL070972 Drosophill
C 28	50.4	7.2	805	12	CNS06V7I	AL416788 T7 end of
C 29	50.4	7.2	1101	12	CNS012TP	AL102007 Drosophill
C 30	50.2	7.2	886	12	BH177277	BH177277 008_L22
C 31	50.2	7.2	886	12	CNS07JUX	AL614235 T3 end of
C 32	50.2	7.2	1101	12	CNS016LI	AL106896 Drosophill
C 33	49.8	7.1	664	12	BH185157	BH185157 036_D_04-
C 34	49.8	7.1	664	12	CNS07PXL	AL622107 T7 end of
C 35	49.8	7.1	928	12	CNS00DKY	AL071865 Drosophill
C 36	49.8	7.1	931	12	CNS005IS	AL059553 Drosophill
C 37	49.6	7.1	1176	12	CNS03R83	AL236908 Tetraodon
C 38	49.4	7.1	1101	12	CNS003BD	AL064091 Drosophill
C 39	49	7.0	949	12	CNS04ATH	AL281906 Tetraodon
C 40	49	7.0	1029	12	CNS012GM	AL174271 Tetraodon
C 41	48.8	7.0	919	12	CNS0058L	AL061409 Tetraodon
C 42	48.8	7.0	945	12	CNS04DOK	AL285149 Tetraodon
C 43	48.8	7.0	1101	12	CNS0039Q	AL063931 Drosophill
C 44	48.6	6.9	306	9	AU037189	AU037189
C 45	48.6	6.9	379	9	AL514359	AL514359

#### ALIGNMENTS

RESULT 1	CNS0039G	1101 bp	DNA	linear	GSS 03-JUN-1999
CNS0039G	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
LOCUS	BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
DEFINITION	AL063921				
ACCESSION	AL063921.1	GI:4941778			
VERSION	GSS.				
KEYWORDS	fruit fly.				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999); Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .				
FEATURES	Location/Qualifiers				
source	1..1101				
	/organism="Drosophila melanogaster"				
	/db_xref="taxon:7227"				
	/clone_lib="RPCI-98"				

```

/clone="BACR08K10"
/note="end : TET3"
BASE COUNT 201 a 64 c 131 g 202 t 503 others
ORIGIN

Query Match 9.2%; Score 64.2; DB 12; Length 1101;
Best Local Similarity 17.6%; Pred. No. 0.51;
Matches 100; Conservative 255; Mismatches 212; Indels 2; Gaps 2;

QY 19 TTTTTCCTTTTACCTATTCAATTAGTTCTTTCATTCAACAGCATGCAAGTCAATTAT 78
Db 483 TTTTTCCTTTTACCTATTCAATTAGTTCTTTCATTCAACAGCATGCAAGTCAATTAT 78
QY 79 TTTTTCCTTTTACCTATTCAATTAGTTCTTTCATTCAACAGCATGCAAGTCAATTAT 138
Db 543 TTTTTCCTTTTACCTATTCAATTAGTTCTTTCATTCAACAGCATGCAAGTCAATTAT 602
QY 139 AATTCAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 198
Db 603 AATTCAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
QY 199 GGAACACGCTCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 258
Db 662 AATTCAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
QY 259 GTTTCGAAAGATGAGACATGCTTCTTAAACCCCTTAAATCGCGCTGAGATGATGAT 318
Db 722 GTTTCGAAAGATGAGACATGCTTCTTAAACCCCTTAAATCGCGCTGAGATGATGAT 780
QY 319 CTTTTCGCTGAGAGAGCTAGGACACATCTCAATCTTCTTCTTCTTCTTCTTCTTCT 378
Db 781 CTTTTCGCTGAGAGAGCTAGGACACATCTCAATCTTCTTCTTCTTCTTCTTCTTCT 840
QY 379 TTTTTCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
Db 841 TTTTTCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 439 ACAAAATTCCTTCTGAGAGCTGCGGAGAACTCAATCTTCTTCTTCTTCTTCTTCTTCT 498
Db 901 ACAAAATTCCTTCTGAGAGCTGCGGAGAACTCAATCTTCTTCTTCTTCTTCTTCTTCT 960
QY 499 TATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 558
Db 961 TATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 559 TGAATAATCAAGATCAAAATATTTCAAAA 587
Db 1021 TGAATAATCAAGATCAAAATATTTCAAAA 1049

RESULT 2
CNS020K7 1092 bp DNA linear GSS 12-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 222111 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION ALI75696
VERSION ALI75696.1 GI:7813753
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.

REFERENCE
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizesma,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
JOURNAL freshwater pufferfish Tetraodon nigroviridis
REFERENCE 2 (bases 1 to 1092)
```

```

AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizesma,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 1092)
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
source 1..1092
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="222111"
/clone_lib="G"
/note="Genoscope sequence ID : COAG222CF06LPI-end : T7"
BASE COUNT 383 a 169 c 165 g 262 t 113 others
ORIGIN

Query Match 8.7%; Score 61; DB 12; Length 1092;
Best Local Similarity 36.8%; Pred. No. 1.6;
Matches 157; Conservative 54; Mismatches 216; Indels 0; Gaps 0;

QY 236 TTTGTTTCTTATACAAAACATGTTTCAAAAAGATGAGACATGCTTAAACCCCTT 295
Db 645 TTTGTTTCTTATACAAAACATGTTTCAAAAAGATGAGACATGCTTAAACCCCTT 704
QY 296 TAATCGCGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 355
Db 705 TAATCGCGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764
QY 356 ATCTTCTTAAATTCGATATATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 415
Db 765 ATCTTCTTAAATTCGATATATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 824
QY 416 TCCAAATTAACACAAATAAATCAACAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCT 475
Db 825 TCCAAATTAACACAAATAAATCAACAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCT 884
QY 476 CTCCCAATTTGGATTTGGATTTGGATTTGGATTTGGATTTGGATTTGGATTTGGATTT 535
Db 885 CTCCCAATTTGGATTTGGATTTGGATTTGGATTTGGATTTGGATTTGGATTTGGATTT 944
QY 536 GCCTATAAAGAAATGCGCCTTGAATTAATCAAGAAATCAAAATATTCAAAATGAAATA 595
Db 945 GCCTATAAAGAAATGCGCCTTGAATTAATCAAGAAATCAAAATATTCAAAATGAAATA 1004
QY 596 TGAAGCCCTAAACTCAACAAATCCAGATCAGGTGGGATTAACCTTAACCTCAGAACCAA 655
Db 1005 TGAAGCCCTAAACTCAACAAATCCAGATCAGGTGGGATTAACCTTAACCTCAGAACCAA 1064
QY 656 ACACAAA 662
Db 1065 ATAAAAA 1071

RESULT 3
CNS020N7 1190 bp DNA linear GSS 14-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
DEFINITION 153P04 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL206908
VERSION AL206908.1 GI:7865727
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.

REFERENCE
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizesma,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
JOURNAL freshwater pufferfish Tetraodon nigroviridis
REFERENCE 2 (bases 1 to 1092)
```

[illegible]

```
Db 418 AAAATTTTAAATTTATATATAAAGGTTATATATATATATAAAGAA-TAAATTTCAATAT 476
QY 542 AAAAGAAATGCGCCACTTGAATAATCAAGAATCAAAATATTCAAATGAATGAAGC 601
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 477 WAAAAGAAATGTTAAAAAATAATTTWAAAAAATAATTTAAATATATATTTTAAAAAT 536
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
QY 602 CACTAAATCAACAATCCAGATCAGTGGGAATA 637
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 537 WMTTATATATTAAAAATTAATTAATANTNNNNNA 572

RESULT 5
CNS00DKY/c
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence T7 end of BAC #
  BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL071865.1 GI:4948170
VERSION
  GSS.
KEYWORDS
  fruit fly.
SOURCE
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
  Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 928)
AUTHORS
  Genoscope.
TITLE
  Direct Submission
JOURNAL
  Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
COMMENT
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the Berkeley Drosophila Genome Project (BDGP).
  The BDGP is constructing a physical map of the Drosophila
  melanogaster genome using these BACs. For further information
  please see http://www.fruitfly.org The BDGP Drosophila
  melanogaster BAC library was prepared by Kazutoyo Osoegawa and
  Aaron Mammosier in Pieter de Jong's laboratory in the Department of
  Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
  NY. The library is named RPCI-98 and was constructed by partial
  EcoRI digestion of Drosophila DNA provided by the BDGP from the
  isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
  P1 and EST libraries. A more detailed description of the library
  and how to order individual BAC clones, the entire library, or
  filters for hybridization from the BACPAC Resource Center can be
  found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
  source
    1..928
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone_lib="RPCI-98"
    /clone="BACR27A24"
    /note="end : T7"
BASE COUNT
  262 a 70 c 84 g 321 t 191 others
ORIGIN
  Query Match 7.8%; Score 54.8; DB 12; Length 928;
  Best Local Similarity 28.9%; Pred. No. 15;
  Matches 92; Conservative 83; Mismatches 143; Indels 0; Gaps 0;

QY 307 AGATAAATGACATTTGGTTCTCAGAAAGACATTAGCGACAACTCTCAATCTCTAATT 366
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 875 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 816
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
QY 367 CTGAATATTCTTTCTTACATGGAATAACTGCATCTTGGATATATTCCAATTAAACA 426
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 815 HHHHTTTTTTTTWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 756
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
QY 427 CAATAAATCAACAATAATTTCTTGAAGTTCGCGAGAAACATCAATCTCCACAATTG 486
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 755 NNCWAWATWWWATATTTATATATATATATATATATATATATATATATATATATAT 696
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
QY 487 GATATTGAGATGATTCATCGATGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 546
```

```
Db 695 TTTATTTTAAATTTTATATATATATATATATATATATATATATATATATATATATAT 636
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
QY 547 AAAATGGCCACTTGAATAATCAAGAATCAAAATATTTCAAAATGAATGAAGCCACTA 606
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 635 TATATATATATATATTTTWTATTTAAATATATATATATATATATATATATATATAT 576
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
QY 607 AACTCAACAATCCAGAA 624
      :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||
Db 575 AAWTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 558

RESULT 6
CNS0100X/c
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence SP6 end of BAC
  BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL098379.1 GI:5609990
VERSION
  GSS.
KEYWORDS
  fruit fly.
SOURCE
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
  Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 1101)
AUTHORS
  Genoscope.
TITLE
  Direct Submission
JOURNAL
  Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
COMMENT
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the European Drosophila Genome Project (EDGP) -
  http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
  library (Dros BAC) was made by Alain Billaud at CEPH (Centre
  d'Etude du Polymorphisme Humain) with funding provided by a MRC
  project grant. The DNA was prepared from embryos by Alain Bucheton
  and Genevieve Payan. It has been constructed in the vector
  pBelobAC11.
FEATURES
  source
    1..1101
    /organism="Drosophila melanogaster"
    /plasmid="pBelobAC11"
    /db_xref="taxon:7227"
    /clone_lib="DrosBAC"
    /clone="BACN03G04"
    /note="end : SP6"
BASE COUNT
  195 a 108 c 131 g 161 t 506 others
ORIGIN
  Query Match 7.8%; Score 54.8; DB 12; Length 1101;
  Best Local Similarity 19.1%; Pred. No. 14;
  Matches 97; Conservative 195; Mismatches 217; Indels 0; Gaps 0;

QY 190 TCCTTGATAGGAACACGCTCTTGAAGCTGTAGCTTATTCATCGCTGTGTTTCTATA 249
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 898 TMAATMYMYMHMYCYCYTMYTCTCYCHYCCYCYMYHMYHMYHMYHMYHMYHMYHMYH 839
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 250 CAAAAACATGTTTGAAGAAGATGAAGACATTCCTTAACACCTTTAATCGCCTCAGA 309
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 838 WATWTACAHHTTMMTMMTMMTMMTMMTMMTMMTMMTMMTMMTMMTMMTMMTMMTMMTMM 779
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 310 ATAAATGACTTTTGGTTCTGAGAAAGACTTAGCGACAACTCTCAATCTCTAATCTG 369
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 778 CHYMTMHMHMYHMYTAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMA 719
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 370 AATAATATCTTTCTTACATGGAATAACTGCATCTTGGATATATTCCAATTAACACA 429
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 718 THTTTTATTTTWTYHCTMMSHCHCCACMHYHCHYHCHYHCHYHCHYHCHYHCHYHCH 659
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 430 ATAAATCAACAATTTCTTCTTGAAGTTCCGAGAAACATCAATCTCCCAATTCGAT 489
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```













<b>Drosophila melanogaster</b>					
genome survey sequence T7 end of BAC BACN11G1 from Drosophila melanogaster (fruit fly), genomic survey sequence.					
AL103735					
GI:5615346					
GSS.					
fruit fly.					
Drosophila melanogaster					
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;					
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.					
(bases 1 to 1043)					
Genoscope.					
Direct Submission					
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr ) Web : www.genoscope.cns.fr) determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBac11.					
Location/Qualifiers					
..1043					
/organism="Drosophila melanogaster"					
/plasmid="pBelOBac11"					
/db_xref="taxon:7227"					
/clone_lib="DrosBAC"					
/clone="BACN11G1"					
/notes=end : T7*					
BASE COUNT    277 a     96 c    121 g    382 t    167 others					
ORIGIN					
Query Match                  7.5% ; Score 52.2; DB 12; Length 1043;					
Best Local Similarity        36.8%; Pred. No. 36;					
Matches 175; Conservative 67; Mismatches 228; Indels 5; Gaps 2;					
QY	224	TTATTCAATCGCTGTGTTTCATAAAAAACAAGCATGGTTTTGAAGAAGAAAGAACACATGC	283		
:	:	:      :	:		
DB	536	YTDTTTTTTTTTCTCCTTTTTTTTATATAWAATMTWTATATATAAAAATAMTATATATAT	595		
:	:	:      :	:		
QY	284	TC TTAAACCCTTTTAATCGCGCTGAGAATAAATCACATTTTGTTCTGAGAAGACTTAGC	343		
:	:	:      :	:		
DB	596	TTTTATATWAANAATAATWATATWTTTTTTWHTWTWTATTTWWHAATAAT--- 652			
QY	344	GACAACTCCCAATCTCTTAATTCGTAATATATATCTTTTCCTTACATGGAATAACTGCAT	403		
:	:	:      :	:		
DB	653	-ATTTTTTTTWTWTWTWTWTATATATATTTATTTTWTWTWATAAAATATTTTTTATATAT	711		
:	:	:      :	:		
QY	404	TC TTGGATATATCCAATTAACACAAATAAATAACACAAAATTTCTCTGAAACTGCCGG	463		
:	:	:      :	:		
DB	712	TTATATWATWTTTWWAATTTATATATTTTNANAT--ANTTTTTTTTTTTTTTTTTTTTT	770		
QY	464	AGAAACTCAAATCCCACAATTTGGATATTCGAGTATATTCGATGGATTTTCTCTCTCTTT	523		
:	:	:      :	:		
DB	771	TTTTTTTTTAA AAA AA	830		
QY	524	AATTTAAACATCGCGCTATATAAGAAAATGGCCACTTTCGAATAATCAAGATCAAAATATTC	583		
:	:	:      :	:		
DB	831	CYYTAWATATATWAWATAWAWAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	890		
QY	584	AAAATGAATTAATGAAGCCACTTAACCTCAACAAATCCAGAAATCAGGTGGGAATAACTTAA	643		
:	:	:      :	:		
DB	891	AAAHAYATACMAAANCRAATHAAWMAYHAAAWATATAAAAAAAAAAAAAAAAAAAAA	950		
QY	644	CTCAGAACCAACACAAAGCCTCATCCAGTTGTTAGACCGCATCTCTACAGAAA	698		
:	:	:      :	:		
DB	951	AAHHMMHMMAAAAAACCCCCCCCCCCCCCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1005		
:	:	:      :	:		

```

RESULT 15
CNS016CO      1200 bp   DNA       linear    GSS 26-JUL-1999
LOCUS         BACN15E04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
DEFINITION    fruit fly, Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyraroidea; Drosophilidae; Drosophila.
ACCESSION     ALI06578 .1 GI:5622626 GSS.
VERSION       ALI06578.1
KEYWORDS      Genoscope. Direct Submission
SOURCE        Submitted (23-JUL-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
ORGANISM      Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster HAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES             Location/Qualifiers
source               1..1200
                    /organism="Drosophila melanogaster"
                    /plasmid="pBelOBAC11"
                    /db_xref="taxon:7227"
                    /clone_lib="DrosBAC"
                    /clone="BACN15E04"
                    /note="end : T7"

BASE COUNT          351 a   188 c   167 g   307 t   187 others
ORIGIN

Query Match              7.5% ; Score 52.2 ; DB 12 ; Length 1200 ;
Best Local Similarity   35.9% ; Pred.No. 33 ;
Matches                98 ; Mismatches 125 ; Indels 0 ; Gaps 0 ;

Qy  15 ATTTTTCCTTTTACCCTTAATCAATTAGTTCTCTTGATTCAAACGATGCNAAGTCA 74
      :|||:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  1081 WTTTTTTTTTTTTWTATWTTTGTCTTTTWTWMAAATWTTWMAATWTTATTTATAATTWA 1022

Qy  75 TTATTGTGTTTAATCCAGATCGGGATTTTGAANAACATTTCTACTAAATTCGAA 134
      :|||:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  1021 ATAATTWTATWTTWTATWTTATWTTATWAAAAAAAWTTTTTWTTTAAAATTTTTCGTTAAWAA 962

Qy  135 ATAAATTTGACGVTGGCTTGATAAATCGTAGTTGGCTCGGAGCTGATATTTTATCCTT 194
      :|||:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  961 ATTAAATTTAAAAAAWAWTTTTTTTRTTWWAAATTAATTTTAAAATTTTTCGTTAAWAA 902

Qy  195 GATAGGAACAACGGCTCTGAAGCTGTAGTTATTCAATCGCTGCTGTTCTATACAAA 254
      :|||:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  901 AWTCHTTATTTTWTARAAAAATTTTAAATWTTTWTWTWTTTWTWTTWTTATWWWAAACCDAAA 842

Qy  255 ACATGTTTGTAAAAAGAATGAAGCACATTTGTCCT 287
      :|||:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  841 AAACTTWAGHAARGCYYYAMCCCCCKKKKKK 809

Search completed: November 2, 2002, 01:47:13
Job time : 1251.45 secs

```

Search completed: November 2, 2002, 01:47:13  
Job time : 1251.45 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:24:39 ; Search time 864.101 Seconds  
(without alignments)  
15087.634 Million cell updates/sec

Title: US-09-881-556-4  
Perfect score: 623  
Sequence: 1 cctcaaatggtgaacatcaa.....aaaaaaaaaaaaaacctcgag 623

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*

2: gb\_hug.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

-----

1	617	99.0	1310	3	AF178459	AF178459 Cryptospor
2	58	9.3	110000	2	PFMAL13P2_0	AL049185 Plasmodiu
3	57.4	9.2	155204	2	AC007926	AC007926 Trypanoso
4	57.2	9.2	1993	8	MGU36923	U36923 Magnaporthe
5	54.8	8.8	126038	2	AC016068	AC016068 Homo sapi
6	54.4	8.7	70332	2	AC044851	AC044851 Homo sapi
7	54	8.7	70411	6	AX344659	AX344659 Sequence
8	54	8.7	117193	2	AC099851	AC099851 Homo sapi
9	53	8.5	12029	3	AE001381	AE001381 Plasmodiu
10	53	8.5	235183	2	AC079427	AC079427 Mus muscu
11	52.6	8.4	5772	3	AF466309	AF466309 Dictyoste
12	52.6	8.4	199551	2	AC006281	AC006281 Plasmodiu
13	52.2	8.4	2145	3	DDISASIA	H34456 D.discoiden
14	52.2	8.4	5191	3	PFMAL1475	PF361475 Dictyoste
15	52.2	8.4	77835	2	PFMAL13P2_3	Continuation (4 of
16	52	8.3	6183	6	AX346099	AX346099 Sequence
17	52	8.3	9539	6	AX277889	AX277889 Sequence
18	52	8.3	9539	6	AX323566	AX323566 Sequence
19	52	8.3	81120	2	AC022851	AC022851 Homo sapi
20	51.8	8.3	6915	6	AX345308	AX345308 Sequence
21	51.8	8.3	64934	2	AL606723	AL606723 Danio rer
22	51.8	8.3	174783	2	AC021317	AC021317 Homo sapi
23	51.8	8.3	192911	2	AC067923	AC067923 Homo sapi
24	51.8	8.3	310779	2	AC005140	AC005140 Plasmodiu
25	51.2	8.2	3706	3	AF126719	AF126719 Plasmodiu
26	51.2	8.2	64789	2	AC083839	AC083839 Homo sapi
27	51.2	8.2	178273	2	AC005308	AC005308 Plasmodiu
28	51	8.2	840	8	CNS0180K	AL110675 Botrytis
29	50.8	8.2	584	3	AF201315	AF201315 Dugesia p
30	50.8	8.2	6494	6	AX346295	AX346295 Sequence
31	50.8	8.2	6494	6	AX348524	AX348524 Sequence
32	50.6	8.1	189893	2	AC013254	AC013254 Drosophil
33	50.4	8.1	256172	2	AC005139	AC005139 Plasmodiu
34	50.2	8.1	652	9	HS329294	AJ329294 Homo sapi
35	50.2	8.1	1757	6	AX347323	AX347323 Sequence
36	50.2	8.1	7306	6	AX346539	AX346539 Sequence
37	50.2	8.1	23016	3	TRBRPGEN	H94286 Trypanosoma
38	50.2	8.1	74412	2	AC020767	AC020767 Homo sapi
39	50.2	8.1	172770	2	AC023133	AC023133 Homo sapi
40	50	8.0	143585	2	AC013349	AC013349 Homo sapi
41	50	8.0	182690	2	AC009998	AC009998 Homo sapi
42	49.8	8.0	14006	6	AX346860	AX346860 Sequence
43	49.8	8.0	91589	8	AC006194	AC006194 Arabidops
44	49.6	8.0	528	3	AF201321	AF201321 Dugesia p
45	49.6	8.0	5181	6	AX348875	AX348875 Sequence

#### ALIGNMENTS

RESULT 1  
AF178459

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

AF178459 1310 bp mRNA linear INV 01-SEP-2000  
Cryptosporidium parvum unknown mRNA sequence.

AF178459

AF178459

AF178459.1 GI:9957266

Cryptosporidium parvum.

Cryptosporidium parvum

Eukaryota: Alveolata; Apicomplexa; Coccidia; Elmeriida;

Cryptosporidiidae; Cryptosporidium.

1 (bases 1 to 1310)

Simonsen,J.N., Kramer,M., Lowden,C. and Wilkins,J.

Neutralising antigen 2 recognized by human immune sera

Unpublished

2 (bases 1 to 1310)

Simonsen,J.N., Kramer,M., Lowden,C. and Wilkins,J.

Direct Submission

Submitted (16-AUG-1999) Medical Microbiology, University of

Manitoba, 730 William Ave, Winnipeg, MB R3E 0W3, Canada

NCBI staff are still waiting for submitters to provide appropriate

feature information.

Location/Qualifiers

source	1..1310	/organism="Cryptosporidium parvum"	
BASE COUNT	470 a	253 c	206 g 381 t
ORIGIN			
Query Match	99.0%;	Score 617;	DB 3; Length 1310;
Best Local Similarity	100.0%;	Pred. No. 1.8e-113;	
Matches 617;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	CCTCAATGGTGAACATCAAGATCTGGTTCAGAGCAAGCCCTATTACTCACCAGAAA	60
Db	694	CCTCAATGGTGAACATCAAGATCTGGTTCAGAGCAAGCCCTATTACTCACCAGAAA	753
QY	61	ACGAATCAAGTTCAAAATCATCTTCTGTGACAGATTCAGATCTGATCAGTTCAAATCT	120
Db	754	ACGAATCAAGTTCAAAATCATCTTCTGTGACAGATTCAGATCTGATCAGTTCAAATCT	813
QY	121	CCTTCTGTACTATTCAGAGACTGGATCAGACTCAGATCAGCGCGCTTGTGACAATCC	180
Db	814	CCTTCTGTACTATTCAGAGACTGGATCAGACTCAGATCAGCGCGCTTGTGACAATCC	873
QY	181	AGAGACTGGATCAGTTCAAATCATCTTCTGCTACTATACAGAAACAGGATCCAGCTCAG	240
Db	874	AGAGACTGGATCAGTTCAAATCATCTTCTGCTACTATACAGAAACAGGATCCAGCTCAG	933
QY	241	ATCACACTCTGTACTTCTCCAGAAGAGATTGGACTCAGAACGTTACCAATCACTTCT	300
Db	934	ATCACACTCTGTACTTCTCCAGAAGAGATTGGACTCAGAACGTTACCAATCACTTCT	993
QY	301	ACAGAACAACCTCAAGCAGCTACATATCTTAACCAAGAAATGAANAATCAATAATC	360
Db	994	ACAGAACAACCTCAAGCAGCTACATATCTTAACCAAGAAATGAANAATCAATAATC	1053
QY	361	AGGAAGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCACTTTCACGCA	420
Db	1054	AGGAAGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCACTTTCACGCA	1113
QY	421	AGCTGCAGATGTGGAAGTTATGGGAAAGAGATAAAATGGTTGATGGTGAGCAAGTAA	480
Db	1114	AGCTGCAGATGTGGAAGTTATGGGAAAGAGATAAAATGGTTGATGGTGAGCAAGTAA	1173
QY	481	TCACATAAATGACATATTGAGATACACTTTCGAAAGAAATAGAAACAAATGTAAGTA	540
Db	1174	TCACATAAATGACATATTGAGATACACTTTCGAAAGAAATAGAAACAAATGTAAGTA	1233
QY	541	CTCGACTTCAATAATATGACCTTAGCCATTTCCTCAATATCTAAATGTCACTCAAGTAA	600
Db	1234	CTCGACTTCAATAATATGACCTTAGCCATTTCCTCAATATCTAAATGTCACTCAAGTAA	1293
QY	601	AAAAAAAAAAAAAAAAAAAA 617	
Db	1294	AAAAAAAAAAAAAAAAAAAA 1310	
RESULT 2			
PFMAL13P2_0			
WPCOMMENT			
Sequence split into 4 fragments	LOCUS	PFMAL13P2	Accession AL049185
Fragment Name	Begin	End	
PFMAL13P2_0	1	110000	
PFMAL13P2_1	100001	210000	
PFMAL13P2_2	200001	310000	
PFMAL13P2_3	300001	377835	
LOCUS	PFMAL13P2	377835 bp	DNA linear HTG 11-AUG-1999
DEFINITION	Plasmodium falciparum chromosome 13 strain 3D7, *** SEQUENCING IN PROGRESS ***, in unordered pieces.		
ACCESSION	AL049185		
VERSION	AL049185.4	GI:5731883	
KEYWORDS	HTG; HTGS_PHASE1.		
SOURCE	malaria parasite P. falciparum.		
ORGANISM	Plasmodium falciparum		
	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
REFERENCE			
AUTHORS			
1 (bases 1 to 377835)			
Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M.			
and Barrell,B.			
Direct Submission			
Submitted (15-MAR-1999)			
P.falciparum Genome Sequencing Consortium,			
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge			
CB10 1SA, UK			
On Aug 12, 1999			
this sequence version replaced gi:5531337.			
For more information about this sequence or the Malaria Project,			
see http://www.sanger.ac.uk/Projects/P_falciparum. IMPORTANT: This			
sequence is unfinished and does not necessarily represent the			
correct sequence. Work on the sequence is in progress and the			
release of this data is based on the understanding that the			
sequence may change as work continues. The sequence may be			
contaminated with foreign sequence from E.coli, yeast, vector,			
phage etc.			
Order of segments is not known; 800 n's separate segments.			
* NOTE: This is a 'working draft' sequence.			
* This record will be updated with the finished sequence			
* as soon as it is available and the accession number will			
* be preserved.			
FEATURES			
Location/Qualifiers			
source			
1..377835			
/organism="Plasmodium falciparum"			
/strain="3D7"			
/db_xref="taxon:5833"			
/chromosome="13"			
BASE COUNT	13224 a	36545 c	35558 g 130270 t 43238 others
ORIGIN			
Query Match	9.3%;	Score 58;	DB 2; Length 110000;
Best Local Similarity	50.4%;	Pred. No. 0.02;	
Matches 142;	Conservative 0;	Mismatches 140;	Indels 0; Gaps 0;
QY	335	CCACAAATGAAATCATATATATCAGAGAGTAAATTCGAGTTTATACACTAAATCT	394
Db	24936	CCACCTTATGAAATCATATATATCAGAGAGTAAATTCGAGTTTATACACTAAATCT	24995
QY	395	TCCAATCAACCCCACTTTCACGCAAGTCGACAGTGGGAAAGCTTATGGGAAAGGA	454
Db	24996	AAAAAAAAGAAAT	25055
QY	455	TAAATGTTGATGGTGAGCAAGTAAATCACTAAAAATGACATTTATGAGATCTCGAA	514
Db	25056	AAAT	25115
QY	515	AGAAATAGAAACAAATGTAAGTATCTGCATTCGATTAATATGCGCTTAGCATTTCCA	574
Db	25116	TAAAAAAAAGAAAT	25175
QY	575	AATATCTAAATTTGCTCACTCAAGTAAAAAAAAGAAATATATATATATATATAT	616
Db	25176	AAAT	25217
RESULT 3			
AC007926/c			
LOCUS	AC007926	155204 bp	DNA linear HTG 17-JUL-2001
DEFINITION	Trypanosoma brucei chromosome II clone RPC193-3H15, *** SEQUENCING IN PROGRESS ***, 1 ordered pieces.		
ACCESSION	AC007926		
VERSION	AC007926.8	GI:14787210	
KEYWORDS	HTG; HTGS_PHASE2.		
SOURCE	Trypanosoma brucei.		
ORGANISM	Trypanosoma brucei		
	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;		
REFERENCE			
AUTHORS			
1 (bases 1 to 155204)			
El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,			
Peterson,J., Hou,L., Zhao,H., Mason,T., Militscher,J., Pai,G., Van			
Aken,S., Utterback,T., Khalak,H.G., Gerard,C., Leech,V., Ollu,E.,			
Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M.			
Trypanosoma brucei GUTat10.1 RPC193-3H15 BAC genomic sequence			

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 155204)  
AUTHORS El-Sayed,N.M., Khalak,H. and Adams,M.D.  
TITLE Direct Submission  
JOURNAL Submitted (28-JUN-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
COMMENT On Jul 17, 2001 this sequence version replaced gi:12746529.  
\* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and the accession number will be preserved.  
\* 1 155204: contig of 155204 bp in length.

FEATURES  
Location/Qualifiers  
1..155204  
/organism="Trypanosoma brucei"  
/isolate="GUTat10.1"  
/db\_xref="taxon:5691"  
/chromosome="II"  
/clone="RPC193-3H15"  
BASE COUNT 48747 a 31603 c 32329 g 42525 t  
ORIGIN

Query Match 9.2% Score 57.4; DB 2; Length 155204;  
Best Local Similarity 47.6% Pred. No. 0.025;  
Matches 169; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

Qy 262 AGAAGAAGGATTGGACTCAGAACGTTACCAATCAGTCTTACAGAACAACTCAAGCCAG 321  
Db 77407 ATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 77348  
Qy 322 CTACATATCCCAACCAAGAAATGAAATCATATAATCAGGAAGTAAATTCGAGTTTAA 381  
Db 77347 ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 77288  
Qy 382 ATACACTAAATCTTCCAAATCAAGCCAACTTTCCAGCAAGCTGGCAGATGTGGAAGTT 441  
Db 77287 ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 77228  
Qy 442 ATGGGAAAGCATAAATGTTGTGAGCAAGTAAATCAGTAAATGACATTAATG 501  
Db 77227 ATGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 77168  
Qy 502 AAGATACATCCAAAGAAATGAAACAAATGTAAGTATCTGCATTCATAAATATGCC 561  
Db 77167 GGATAGTGTGTGAGTGTATATACGAATAATAATAAGAGCAGTAAATAATAACA 77108  
Qy 562 TTAGCCATTTCCAAATATCTAANTTGTCAACTCAAGTAAAAAAGAGAGAGAGAGAG 616  
Db 77107 ATTGAATTTATCAGAGAAATGATACAAATGAAAGAGAGAGAGAGAGAGAGAGAG 77053

RESULT 4  
MGU36923  
LOCUS Magnaporthe grisea host-species specificity (Pw1) gene, complete cds.  
DEFINITION MGU36923 1993 bp DNA linear PLN 13-DEC-2001  
ACCESSION U36923  
VERSION U36923.1 GI:1045530  
KEYWORDS Magnaporthe grisea.  
SOURCE Magnaporthe grisea.  
ORGANISM Magnaporthe grisea.  
Eukaryote; Fungi; Ascomycota; Perizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.  
REFERENCE 1 (bases 1 to 1993)  
AUTHORS Kang,S., Sweigard,J.A. and Valent,B.  
TITLE The Pw1 host specificity gene family in the blast fungus Magnaporthe grisea  
JOURNAL Mol. Plant Microbe Interact. 8 (6), 939-948 (1995)

MEDLINE 96112748  
PUBMED 8664503  
REFERENCE 2 (bases 1 to 1993)  
AUTHORS Kang,S., Sweigard,J.A. and Valent,B.  
TITLE Direct Submission  
JOURNAL Submitted (22-SEP-1995) Seogchan Kang, Biology, University of New Mexico, Albuquerque, NM 87131, USA  
COMMENT Sweigard, J.A. Plant Cell 7, 1221-1233, 1995.  
FEATURES  
Location/Qualifiers  
1..1993  
/organism="Magnaporthe grisea"  
/db\_xref="taxon:148305"  
/gene="Pw1"  
/note="Pw1 host species specificity gene"  
/codon\_start=1  
/product="Pw1 protein"  
/protein\_id="AAA80239.1"  
/db\_xref="GI:1045531"  
/translation="MLFTALVNLPAVRVRRIDSGAPMISFNIALLLIHLHYCRFNKK  
FNKTIPLIILAFSTAVIAGRKWTKVYNDKGPREGSISIRKGAEGDNCGPYGP  
GPDMMRVVHEDNGNIRGMPGYPYRLGDDKEDKDNQYYSRNGYHVGDPAPYQNHGG  
OMGDGYGPPGQITNQHGRQDGGCHM"  
BASE COUNT 778 a 313 c 372 g 530 t  
ORIGIN

Query Match 9.2% Score 57.2; DB 8; Length 1993;  
Best Local Similarity 49.0% Pred. No. 0.062;  
Matches 152; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 308 AAATCTCAAGCAGCTACATATCTTACCAATCAACCCCAATCTTCCACCAAGCTGGC 427  
Db 1479 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1538  
Qy 368 TAATTCGAGTTTATACACATAATCTTCCAAATCAACCCCAATCTTCCACCAAGCTGGC 427  
Db 1539 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1598  
Qy 428 AGATGTGMAAGTTATGGGAAAAGGATATAATGTTGATGGTGAAGCAAGTAAATCACTAA 487  
Db 1599 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1658  
Qy 488 AAATGACATTAATGAAGATACCTCGAAGAAATAGAAACAAATGTAAAGTATCTGCAT 547  
Db 1659 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1718  
Qy 548 TGATAAATATGGCTTAGCCATTTCCCAATATCTAAATTTGCAACTCAAGTAAAAAAGAG 607  
Db 1719 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1778  
Qy 608 AAAAAAANA 617  
Db 1779 AAAAAAANA 1788

RESULT 5  
AC016068  
LOCUS Homo sapiens chromosome 4 clone RP11-431B5 map 4, LOW-PASS SEQUENCE SAMPLING.  
DEFINITION AC016068  
ACCESSION AC016068  
VERSION AC016068.3 GI:9126605  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 126038)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens chromosome 4, clone RP11-431B5  
JOURNAL Unpublished

REFERENCE  
AUTHORS

2 (bases 1 to 126038)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Castle,A., Cangelino,M., Collins,S., Collymore,A.,  
 Cooke,P., DeAtellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,  
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (20-NOV-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 13, 2000 this sequence version replaced 91:6715901.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE  
JOURNAL

## COMMENT

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L4278  
 Center clone name: 431\_B\_5  
 -----

\* NOTE: This record contains 130 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1  
 \* 890 989: contig of 889 bp in length  
 \* 990 1894: contig of 905 bp in length  
 \* 1895 1994: gap of 100 bp  
 \* 1995 2876: contig of 882 bp in length  
 \* 2877 2976: gap of 100 bp  
 \* 2977 3868: contig of 892 bp in length  
 \* 3869 3968: gap of 100 bp  
 \* 3969 4868: contig of 900 bp in length  
 \* 4869 4968: gap of 100 bp  
 \* 4969 5882: contig of 914 bp in length  
 \* 5883 5982: gap of 100 bp  
 \* 5983 6879: contig of 897 bp in length  
 \* 6880 6979: gap of 100 bp  
 \* 6980 7908: contig of 929 bp in length  
 \* 7909 8008: gap of 100 bp  
 \* 8009 8908: contig of 900 bp in length  
 \* 8909 9008: gap of 100 bp  
 \* 9009 9870: contig of 862 bp in length  
 \* 9871 9970: gap of 100 bp  
 \* 9971 10879: contig of 909 bp in length  
 \* 10880 10979: gap of 100 bp  
 \* 10980 11863: contig of 884 bp in length  
 \* 11864 11963: gap of 100 bp  
 \* 11964 12838: contig of 875 bp in length  
 \* 12839 12938: gap of 100 bp  
 \* 12939 13829: contig of 891 bp in length  
 \* 13830 13929: gap of 100 bp  
 \* 13930 14854: contig of 925 bp in length  
 \* 14855 14954: gap of 100 bp  
 \* 14955 15863: contig of 909 bp in length

\* 15864 15963: gap of 100 bp  
 \* 15964 16850: contig of 887 bp in length  
 \* 16851 16950: gap of 100 bp  
 \* 16951 17871: contig of 921 bp in length  
 \* 17872 17971: gap of 100 bp  
 \* 17972 18858: contig of 887 bp in length  
 \* 18859 18958: gap of 100 bp  
 \* 18959 19690: contig of 732 bp in length  
 \* 19691 19790: gap of 100 bp  
 \* 19791 20687: contig of 897 bp in length  
 \* 20688 20787: gap of 100 bp  
 \* 20788 21671: contig of 884 bp in length  
 \* 21672 21771: gap of 100 bp  
 \* 21772 22673: contig of 902 bp in length  
 \* 22674 22773: gap of 100 bp  
 \* 22774 23665: contig of 892 bp in length  
 \* 23666 23765: gap of 100 bp  
 \* 23766 24662: contig of 897 bp in length  
 \* 24663 24762: gap of 100 bp  
 \* 24763 25690: contig of 928 bp in length  
 \* 25691 25790: gap of 100 bp  
 \* 25791 26696: contig of 906 bp in length  
 \* 26697 26796: gap of 100 bp  
 \* 26797 27710: contig of 914 bp in length  
 \* 27711 27810: gap of 100 bp  
 \* 27811 28693: contig of 883 bp in length  
 \* 28694 28793: gap of 100 bp  
 \* 28794 29700: contig of 907 bp in length  
 \* 29701 29800: gap of 100 bp  
 \* 29801 30699: contig of 899 bp in length  
 \* 30700 30799: gap of 100 bp  
 \* 30800 31699: contig of 900 bp in length  
 \* 31700 31799: gap of 100 bp  
 \* 31800 32718: contig of 919 bp in length  
 \* 32719 32818: gap of 100 bp  
 \* 32819 33744: contig of 926 bp in length  
 \* 33745 33844: gap of 100 bp  
 \* 33845 34734: contig of 890 bp in length  
 \* 34735 34834: gap of 100 bp  
 \* 34835 35771: contig of 937 bp in length  
 \* 35772 35871: gap of 100 bp  
 \* 35872 36789: contig of 918 bp in length  
 \* 36790 36889: gap of 100 bp  
 \* 36890 37797: contig of 908 bp in length  
 \* 37798 37897: gap of 100 bp  
 \* 37898 38819: contig of 922 bp in length  
 \* 38820 38919: gap of 100 bp  
 \* 38920 39854: contig of 935 bp in length  
 \* 39855 39954: gap of 100 bp  
 \* 39955 40860: contig of 906 bp in length  
 \* 40861 40960: gap of 100 bp  
 \* 40961 41848: contig of 888 bp in length  
 \* 41849 41948: gap of 100 bp  
 \* 41949 42837: contig of 889 bp in length  
 \* 42838 42937: gap of 100 bp  
 \* 42938 43842: contig of 905 bp in length  
 \* 43843 43942: gap of 100 bp  
 \* 43943 44850: contig of 908 bp in length  
 \* 44851 44950: gap of 100 bp  
 \* 44951 45883: contig of 933 bp in length  
 \* 45884 45983: gap of 100 bp  
 \* 45984 46879: contig of 896 bp in length  
 \* 46880 46979: gap of 100 bp  
 \* 46980 47858: contig of 879 bp in length  
 \* 47859 47958: gap of 100 bp  
 \* 47959 48839: contig of 881 bp in length  
 \* 48840 48939: gap of 100 bp  
 \* 48940 49851: contig of 912 bp in length  
 \* 49852 49951: gap of 100 bp  
 \* 49952 50886: contig of 935 bp in length  
 \* 50887 50986: gap of 100 bp  
 \* 50987 51916: contig of 930 bp in length  
 \* 51917 52016: gap of 100 bp

```
* 52017 52912: contig of 896 bp in length
* 52913 53012: gap of 100 bp
* 53013 53911: contig of 899 bp in length
* 53912 54011: gap of 100 bp
* 54012 54890: contig of 879 bp in length
* 54891 54990: gap of 100 bp
* 54991 55881: contig of 891 bp in length
* 55882 55981: gap of 100 bp
* 55982 56858: contig of 877 bp in length
* 56859 56958: gap of 100 bp
* 56959 57860: contig of 902 bp in length
* 57861 57960: gap of 100 bp
* 57961 58859: contig of 899 bp in length
* 58860 58959: gap of 100 bp
* 58960 59855: contig of 896 bp in length
* 59856 59955: gap of 100 bp
* 59956 60872: contig of 917 bp in length
* 60873 60972: gap of 100 bp
* 60973 61865: contig of 893 bp in length
* 61866 61965: gap of 100 bp
* 61966 62867: contig of 902 bp in length
* 62868 62967: gap of 100 bp
* 62968 63861: contig of 894 bp in length
* 63862 63961: gap of 100 bp
* 63962 64858: contig of 897 bp in length
* 64859 64958: gap of 100 bp
* 64959 65851: contig of 893 bp in length
* 65852 65951: gap of 100 bp
* 65952 66849: contig of 898 bp in length
* 66850 66949: gap of 100 bp
* 66950 67838: contig of 889 bp in length
* 67839 67938: gap of 100 bp
* 67939 68873: contig of 935 bp in length
* 68874 68973: gap of 100 bp
* 68974 69863: contig of 910 bp in length
* 69864 69963: gap of 100 bp
* 69964 70890: contig of 907 bp in length
* 70891 70990: gap of 100 bp
* 70991 71930: contig of 940 bp in length
* 71931 72030: gap of 100 bp

Query Match      8.8%  Score 54.8:  DB 2:  Length 126038;
Best Local Similarity 38.5%:  Pred. No. 0.085;
Matches 137:  Conservative 0:  Mismatches 219:  Indels 0:  Gaps 0:

Qy 262 AGAAGAAGGATTGGACCTCAGAACGTTACCAATCACTTCTACAGAACAACTCAAGCCAG 321
Db 122722 ANNAANNAANNAANNAANNAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 122781
Qy 322 CTACATATCTTACCAAGAAATGAAATCATATATATATATATATATATATATATATAT 381
Db 122782 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 122841
Qy 382 ATACACTAAATCTTCCAAATCAACCCCAATCTTTCACGCAAGCTGGCAGATCTGCAAGTT 441
Db 122842 ANNAANNAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 122901
Qy 442 ATGCGGAAAGGATATAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 501
Db 122902 ANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 122961
Qy 502 AAGATATCTTGAAGAAATTTAGAACCAAAATGTTAAAGTATCTGCAATGATATATGCCC 561
Db 122962 ANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 123021
Qy 562 TTACCCATCTCCAAATATCTAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 617
Db 123022 ANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 123077

RESULT 6
AC044851
LOCUS
```

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens chromosome 2 clone RP11-382G9 map 2, LOW-PASS SEQUENCE SAMPLING.  
AC044851  
GI:7543818  
HTG: HTGS\_PHASE0.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 70332)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 2, clone RP11-382G9  
Unpublished  
2 (bases 1 to 70332)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F., Boguslavsky, L., Bouckgealter, B., Brown, A., Burckett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferrelira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRocque, K., Lamazara, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneu, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, T., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tessaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

Direct Submission  
Submitted (12-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L9217  
Center clone name: 382G\_9  
-----

NOTE: This record contains 81 individual  
sequencing reads that have not been assembled into  
contigs. Runs of N are used to separate the reads  
and the order in which they appear is completely  
arbitrary. Low-pass sequence sampling is useful for  
identifying clones that may be gene-rich and allows  
overlap relationships among clones to be deduced.  
However, it should not be assumed that this clone  
will be sequenced to completion. In the event that  
the record is updated, the accession number will  
be preserved.  
1 666: contig of 666 bp in length  
667 766: gap of 100 bp  
767 1534: contig of 768 bp in length  
1535 1634: gap of 100 bp  
1635 2392: contig of 758 bp in length  
2393 2492: gap of 100 bp  
2493 3249: contig of 757 bp in length  
3250 3349: gap of 100 bp  
3350 4123: contig of 774 bp in length  
4124 4223: gap of 100 bp  
4224 4982: contig of 759 bp in length  
4983 5082: gap of 100 bp







```
* 11374 11473: gap of 100 bp
* 11474 12187: contig of 714 bp in length
* 12188 12287: gap of 100 bp
* 12288 12985: contig of 698 bp in length
* 12986 13085: gap of 100 bp
* 13086 13817: contig of 732 bp in length
* 13818 13917: gap of 100 bp
* 13918 14645: contig of 728 bp in length
* 14646 14745: gap of 100 bp
* 14746 15479: contig of 734 bp in length
* 15480 15579: gap of 100 bp
* 15580 16316: contig of 737 bp in length
* 16317 16416: gap of 100 bp
* 16417 17109: contig of 693 bp in length
* 17110 17209: gap of 100 bp
* 17210 17945: contig of 736 bp in length
* 17946 18045: gap of 100 bp
* 18046 18759: contig of 714 bp in length
* 18760 18859: gap of 100 bp
* 18860 19646: contig of 787 bp in length
* 19647 19746: gap of 100 bp
* 19747 20464: contig of 718 bp in length
* 20465 20564: gap of 100 bp
* 20565 21295: contig of 731 bp in length
* 21296 21395: gap of 100 bp
* 21396 22131: contig of 736 bp in length
* 22132 22311: gap of 100 bp
* 22312 22937: contig of 706 bp in length
* 22938 23037: gap of 100 bp
* 23038 23767: contig of 730 bp in length
* 23768 23867: gap of 100 bp
* 23868 24587: contig of 720 bp in length
* 24588 24687: gap of 100 bp
* 24688 25418: contig of 731 bp in length
* 25419 25518: gap of 100 bp
* 25519 26243: contig of 725 bp in length
* 26244 26343: gap of 100 bp
* 26344 27079: contig of 736 bp in length
* 27080 27179: gap of 100 bp
* 27180 27907: contig of 728 bp in length
* 27908 28007: gap of 100 bp
* 28008 28759: contig of 752 bp in length
* 28760 28859: gap of 100 bp
* 28860 29585: contig of 726 bp in length
* 29586 29685: gap of 100 bp
* 29686 30410: contig of 725 bp in length
* 30411 30510: gap of 100 bp
* 30511 31240: contig of 730 bp in length
* 31241 31340: gap of 100 bp
* 31341 32079: contig of 739 bp in length
* 32080 32179: gap of 100 bp
* 32180 32959: contig of 780 bp in length
* 32960 33059: gap of 100 bp
* 33060 33782: contig of 723 bp in length
* 33783 33882: gap of 100 bp
* 33883 34604: contig of 722 bp in length
* 34605 34704: gap of 100 bp
* 34705 35428: contig of 724 bp in length
* 35429 35528: gap of 100 bp
* 35529 36256: contig of 728 bp in length
* 36257 36356: gap of 100 bp
* 36357 37084: contig of 728 bp in length
* 37085 37184: gap of 100 bp
* 37185 37919: contig of 735 bp in length
* 37920 38019: gap of 100 bp
* 38020 38744: contig of 725 bp in length
* 38745 38844: gap of 100 bp
* 38845 39567: contig of 723 bp in length
* 39568 39667: gap of 100 bp
* 39668 40373: contig of 706 bp in length
* 40374 40473: gap of 100 bp
* 40474 41192: contig of 719 bp in length
* 41193 41292: gap of 100 bp
```

```
* 41293 42014: contig of 722 bp in length
* 42015 42114: gap of 100 bp
* 42115 42884: contig of 770 bp in length
* 42885 42984: gap of 100 bp
* 42985 43733: contig of 749 bp in length
* 43734 43833: gap of 100 bp
* 43834 44573: contig of 740 bp in length
* 44574 44673: gap of 100 bp
* 44674 45417: contig of 744 bp in length
* 45418 45517: gap of 100 bp
* 45518 46228: contig of 711 bp in length
* 46229 46328: gap of 100 bp
* 46329 47038: contig of 710 bp in length
* 47039 47138: gap of 100 bp
* 47139 47828: contig of 690 bp in length
* 47829 47928: gap of 100 bp
* 47929 48648: contig of 720 bp in length
* 48649 48748: gap of 100 bp
* 48749 49459: contig of 711 bp in length
* 49460 49559: gap of 100 bp
* 49560 50304: contig of 745 bp in length
* 50305 50404: gap of 100 bp
* 50405 51121: contig of 717 bp in length
* 51122 51221: gap of 100 bp
* 51222 51943: contig of 722 bp in length
* 51944 52043: gap of 100 bp
* 52044 52776: contig of 733 bp in length
* 52777 52876: gap of 100 bp
* 52877 53537: contig of 661 bp in length
* 53538 53637: gap of 100 bp
* 53638 54322: contig of 685 bp in length
* 54323 54422: gap of 100 bp
* 54423 55118: contig of 696 bp in length
* 55119 55218: gap of 100 bp
* 55219 55931: contig of 713 bp in length
* 55932 56031: gap of 100 bp
* 56032 56745: contig of 714 bp in length

Query Match      8.7%; Score 54; DB 2; Length 117193;
Best Local Similarity 27.9%; Pred. No. 0.12;
Matches 126; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 164 CCCCTTGACAAATTCACAGACTGGATCAGTTCAAATCATCTTCTGCTACTATACCAGA 223
DB 62451 CGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 62392
QY 224 RACAGATCCAGCTCAGATCAGCTCTGCTACTTCCAGAGAAGGATTGGACTCAGAA 283
DB 62391 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 62332
QY 284 CGTTACCAATCACTTCTACAGAACAACTCAAAGCCAGCTACATATCTCAACCAAGAAAA 343
DB 62331 ANAANANANANANANANANANANANANANANANANANANANANANANANANANAN 62272
QY 344 TGAATATCATATAATCAGGAAGGTAATTCGAGTTTTTAATACACTAAATCTTCCAAATCA 403
DB 62271 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 62212
QY 404 ACCCAATCTTTCACCAAGCTGGCAGATGTGGAAGTTATGGGGAAGGATAAATGTT 463
DB 62211 NNNNNANANANANANANANANANANANANANANANANANANANANANANANANNA 62152
QY 464 TGATGGTGAGCAAGTAATCACTAAATAATGACATTATTGAAGATACTTCGAAGAAATTTAG 523
DB 62151 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 62092
QY 524 AAACAAATATGTAAGTATCTGCAATGTAATATATGGCCATTTCGCCAATTCCTAAATCTAA 583
DB 62091 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 62032
QY 584 ATTGCTCAACTCAAGTAAAAAANANANANANANANANANANANANANANANANANANAN 615
DB 62031 ANAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 62000
```



## RESULT 10

AC079427  
LOCUS AC079427 235183 bp DNA linear HTG 01-SEP-2000  
DEFINITION Mus musculus chromosome 16 clone RP23-121D10, WORKING DRAFT  
SEQUENCE, 50 unordered pieces.  
AC079427  
AC079427.1 GI:958039  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 235183)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Mouse  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 235183)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (01-SEP-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT -----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 0  
Center clone name: RPCI-23\_121D10  
-----  
Summary Statistics  
Consensus quality: 155133 bases at least Q40  
Consensus quality: 177576 bases at least Q30  
Consensus quality: 186200 bases at least Q20  
Estimated insert size: 193300; agarose-fp estimation  
Estimated insert size: 193283; sum-of-contigs estimation  
Quality coverage: 3.62 in Q20 bases; agarose-fp estimation  
Quality coverage: 3.04 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 50 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1  
1408: contig of 1408 bp in length  
1409: gap of unknown length  
1509: contig of 1542 bp in length  
3051: gap of unknown length  
3151: contig of 1142 bp in length  
4292: gap of unknown length  
4392: gap of 1197 bp in length  
5589: contig of 1197 bp in length  
5689: gap of unknown length  
5690: contig of 1144 bp in length  
6834: gap of unknown length  
6934: contig of 1480 bp in length  
8414: gap of unknown length  
8514: contig of 1229 bp in length  
9742: gap of unknown length  
9842: contig of 1157 bp in length  
9843: gap of unknown length  
10999: contig of 1279 bp in length  
11000: gap of unknown length  
12379: contig of 1125 bp in length  
12479: gap of unknown length  
13603: contig of 1125 bp in length  
13703: gap of unknown length  
13704: contig of 1847 bp in length  
15551: gap of unknown length  
15650: contig of 1739 bp in length  
15651: gap of unknown length  
17390: contig of 1196 bp in length  
17489: gap of unknown length  
17990: contig of 1196 bp in length  
18685: gap of unknown length  
18686: contig of 1442 bp in length  
20227: contig of 1442 bp in length  
20228: gap of unknown length  
21739: contig of 1412 bp in length  
21839: gap of unknown length  
21840: contig of 1337 bp in length  
23177: gap of unknown length  
23277: contig of 1106 bp in length  
24382: gap of unknown length  
24482: contig of 1180 bp in length  
25663: gap of unknown length  
25762: contig of 1734 bp in length  
27496: gap of unknown length  
27597: contig of 1250 bp in length  
28446: gap of unknown length  
28947: contig of 1131 bp in length  
30078: gap of unknown length  
30177: contig of 1143 bp in length  
31321: gap of unknown length  
31420: contig of 1145 bp in length  
32565: gap of unknown length  
32666: contig of 1439 bp in length  
34104: gap of unknown length  
34204: contig of 1023 bp in length  
35227: gap of unknown length  
35328: contig of 1015 bp in length  
36342: gap of unknown length  
36442: contig of 1057 bp in length  
37499: gap of unknown length  
37599: contig of 1264 bp in length  
38863: gap of unknown length  
38964: contig of 1967 bp in length  
40930: gap of unknown length  
41030: contig of 1703 bp in length  
42733: gap of unknown length  
42834: contig of 3783 bp in length  
46616: gap of unknown length  
46716: contig of 3019 bp in length  
49735: gap of unknown length  
49835: contig of 4335 bp in length  
54170: gap of unknown length  
54270: contig of 3212 bp in length  
57482: gap of unknown length  
57582: contig of 5325 bp in length  
62907: gap of unknown length  
63007: contig of 4237 bp in length  
67244: gap of unknown length  
67345: contig of 3510 bp in length  
70854: gap of unknown length  
70855: contig of 8574 bp in length  
70955: gap of unknown length  
79528: contig of 6018 bp in length  
79629: gap of unknown length  
85647: contig of 5389 bp in length  
91135: gap of unknown length  
91235: contig of 9006 bp in length  
100241: gap of unknown length  
100341: contig of 8128 bp in length  
108469: gap of unknown length  
108569: contig of 12974 bp in length  
121543: gap of unknown length  
121643: contig of 11348 bp in length  
132991: gap of unknown length  
133091: contig of 13206 bp in length  
146297: gap of unknown length  
146397: contig of 15570 bp in length  
161967: gap of unknown length  
162067: contig of 14843 bp in length  
176910: gap of unknown length  
177010: contig of 15693 bp in length  
192703: gap of unknown length  
192803: contig of 20276 bp in length  
213079: gap of unknown length  
213179: contig of 22004 bp in length.  
235183: Location/Qualifiers

FEATURES

```

source      1. .235183
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
/clone="RP23-121D10"
/clone_1lb="rPCI mouse BAC library 23"
BASE COUNT 71346 a 43034 c 50925 g 59494 t 10384 others
ORIGIN
Query Match      8.5%; Score 53; DB 2; Length 235183;
Best Local Similarity 39.7%; Pred. No. 0.17;
Matches 140; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

QY 265 AGAAGGATTGGACTCAGAACGTTACCAATCACTTCTACAGAACAACTCAAGCCAGCTA 324
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13942 AAAGGGGNNNNAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAANN 14001

QY 325 CATATCTAACCAAGAAAAATGCAATATCATATAATCAGGAAGTAAATTCGAGTTTTAATA 384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 14002 NNNNNNNNNNAAANAAANNNAAAAAANAAAAAANAAAAAAGGAGGAGAGAGNAAAA 14061

QY 385 CACTAAATCTTCCAAATCAACCCCAATCTTTCACGCAAGCTGGCAGATGTGAAAGTTATG 444
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 14062 GAACAAAAAANNNAAAAAANGNAGAAANAAAGAGAGAGAAAAAAGAAAAAAG 14121

QY 445 GGGAAAGGATAAATCGTTGATCGTGAGCAAGTAAATCACTAAAAATGACATTTATTGAAG 504
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 14122 AAAAAAGAAAGAGAAAGNNGGAAAAAANAAAAAAGAAAAAAGAAAAAGGAAAAAGAGG 14181

QY 505 ATACTTCCAAAGAAATTAGAAAAAATGTAAAGTATCTGCTATTCATAAATATGCCCCTTA 564
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 14182 AGAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 14241

QY 565 GCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAAAAAANAAAAAANAAAAA 617
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 14242 AGAAAGAGAAAAAANAGAAAGAAAAAGGAAAAAAGAAAAAAGGAAAAAGGAAAAAGAAA 14294

```

```

RESULT 11
LOCUS      AF466309
DEFINITION Dictyostellium discoideum serine protease/ABC transporter TagD
            5772 bp DNA linear INV 05-FEB-2002
ACCESSION  AF466309
VERSION    AF466309
KEYWORDS   (tagD) gene, complete cds.
SOURCE     AF466309.1 GI:18496823
            Dictyostellium discoideum.
REFERENCE  Dictyostellium discoideum.
            Eukaryotes; Mycetozoa; Dictyostellida; Dictyostellium.
            1 (bases 1 to 5772)
            Anjard,C. and Loomis,W.F.
            Evolution of the ABC transporters of Dictyostellium
            Unpublished
            2 (bases 1 to 5772)
            Anjard,C. and Loomis,W.F.
            Direct Submission
            Submitted (08-JAN-2002) Biology Dept 0368, UCSD, 9500 Gilman Drive,
            La Jolla, CA 92093, USA
FEATURES   Location/Qualifiers
            1..5772
               /organism="Dictyostellium discoideum"
               /strain="Ax4"
               /db_xref="taxon:44689"
               <243..>5720
               /gene="tagD"
               /product="serine protease/ABC transporter TagD"
               <243..>5720
               /gene="tagD"
               243..5720
               /gene="tagD"
               /note="similar to TagB and TagC"
               /codon_start=1
               /product="serine protease/ABC transporter TagD"

```

/protein\_id="AAU74253.1"  
/db\_xref="GI:18496824"  
/translation="MKSNTNIRVLVSGILIFLFIKPEFINKNNDKIGINRKLKLE  
FSTFKNNNNNNNQDQLKKDKRKLKNEIDTNIKKIKNNKNNDKNEIEFP  
NFTSRLLKNDNMEIQOFTYRKSHYIVQRFKLDINDETRQOKFLINTDVLDEOQY  
SHLVNYIPDSFLVLANDQSNLLSSKEWYSWIGEPESNKIHLNYSKSGIGLPVYIK  
KLVYNSLQWRENTNLSLTYSNKSQVLTILNQKKLKSYVCDNDSOSSOSSGSLVSS  
ESUVYNSIQWSESNTYERSEKFTQANRULSPKAIPTGKTDLVNNDRIDIILPRLCKQIL  
SIATDGLDGHCFSDSNIPYNSVNLNHRKVTYIGSLHDNEDYVDGHGTIVCGSAS  
AGAPEDSLASISFSGLATDAKTAFFDLASDPNNPEPVEDYSQLQYVLNAGRVIVS  
GDSWGSLSIQGYSDDASIDDFLYTHPDFILRAAGNNEQYSLLSSOATAKNYIL  
TVCAGDTHETSYTDALYESNFEFVAKSTLNSLCQSPDDKCYCTYTAQCTTESTVAG  
LSGCTSYIKNSYASTESPESOPELYNENNTSPSKGPTHDCRLKPDIVAGQVITSRAR  
SNQANTDDCGSLNPNTALLSESQTSMATPLATAATTILRQYLDVGYPTGIVGSE  
NKLPQTGSLKALMINNAULNGTTPLSSTNTPSNAAVFDTFAGANVQMGSLRMTSE  
WLXVSEGVKPKPSVSGVIGELGDKKASKNWEYSLSTQNVSYSCFYTPSPSSGSGSG  
GIRIVATLVTDPPSPVSGAKLVNLVNDLTMTNTESEFIFYSNSSGYSVNGTKGTTLL  
NIPNINVEGILYTPINTKSELSFRPLIAGTNPIGPQNFSPVPHGENCEFOWADSC  
MQCPDDTPCPTENGVSQTCDDLYLWGRCLVQSCNNNYNYSIDSKCKFLSYNYIY  
VILVAGTQMLTIVLLIKHMETKMGKFLSUKFQSPGVLGTGCKNVGSGKGGSGG  
SGSGTGLQIDDDGTGHVRPKDAPVTPDPLYSLLSPFIETITSTACSLVATVAA  
SILQYPIGQIDQIPIQDGLRQDFIIFLLALEFVFTISSWIGSEVTKWYVMM  
RLQKVFYRLIAQMGDFQKNSAALVNLVLDTPMLRSLLTGLTSLVSGVICKFVGS  
LVFTIFSNKLSLAFTAPVLAIVTQVQSKFTKRLTRRLLFHNSKASOHGOSVNNK  
HYVSNYCKQDRELAKYSEQLMVPYISRRLLINNTFAASIKMLWVESLAFILLYFGAY  
LAIQOFTVGLVYSFLYIGYIDSSSTLGRVYSYVQCLASATRVFLILRSAPKRRK  
TLSEELNDITDNTQDNNNNNDIDSSDDDDNNNNNNKNNKNNKNNKSSGSDSS  
EADYKKNNKRNNGKMTKLSNSPLVGEIGIDNNNNNNNDINIIDDNQDPNNNNNN  
EIDDQDDGDDGDDGDEGDEGDEGDEGDEGDEGDEGDEGDEGDEGDEGDEGDEG  
GISCELNLTIPSAVTELTCEKGETEFKNFSYPSHADVGVLVNLIDLKFSGKCYGLV  
GPSCKSLTELLISPYLSLHPSGKLYMDGIDIAKIRPSNLBFTNVAHQHPLFVA  
TISENTGYALDNTQEDIDIAKALNAHETIQSLPKQYDTHLTDGGLNSGCKKRIYV  
ALDCAKRRKLMLDEIDPESEAINKSIVLTGHTVYVMVAHVAAVRDCDKIFV  
VLDKGQIEVGTEINQLMAKKGKGYRMAFSEDDDDYAPLLVL"

2189 a 708 c 904 g 1971 t

BASE COUNT  
ORIGIN

Query Match	8.4%	Score 52.6	DB 3	Length 5772
Best Local Similarity	51.5%	Pred. No. 0.42		
Matches 121; Conservative	0	Mismatches 114	Indels 0	Gaps 0
Qy 303	AGAAACAACCTCAAAGCAGCTACATATCCTAACCAAGAAAATGAAAAATCATATAATCAG	362		
Db 4605	ACAAAACATATCAAAATTCACCACCATTAGTAGGTGAAGGTATAGATAATAATAATAAT	4664		
Qy 363	GAAGGTAAATTCGAGTTTTAATACACTAATAATCTTCCAANAATCAACCCAAATCTTTACGGCAAG	422		
Db 4665	AATAATGATAATAATAATAATGCATATAATATCAGCAAGATCCAAAATAATAATAATAAT	4724		
Qy 423	CTGCAGAGTGGAAAGTTATGGGGAAAGCAATAAATGGTTGATGGTGACCAAGTAATC	482		
Db 4725	GAATTTGATGATGATGGTGATGATGATGGTGATGATGATGATGATGATGATGATGATGAT	4784		
Qy 483	ACTAAAAATGACATTAATTGAAGATACTTCGAAAGAAAATTAGAAACAACAAATGTAAA	537		
Db 4785	AATAATAATAATAATAATAATGATGATCCAAATGCATATAATATGGAATTTGAAATGTTAA	4839		

RESULT 12	AC006281	199551 bp	DNA	linear	HTG 12-AUG-2000
LOCUS	AC006281				
DEFINITION	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.				
ACCESSION	AC006281				
VERSION	AC006281.8	GI:9797738			
KEYWORDS	HTG; HTGS, PHASE1.				
SOURCE	malaria parasite P. falciparum.				
ORGANISM	Plasmodium falciparum.				
REFERENCE	Eukaryote; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				
AUTHORS	1 (bases 1 to 199551) Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T., Kurtz, O.B., Conway, A.B. and Davis, R.W.				
TITLE	Plasmodium falciparum 3D7 chromosome 12				

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 199551)  
AUTHORS Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi,O.B. and Davis,R.W.  
TITLE Direct Submission  
JOURNAL Submitted (05-JAN-1999) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

On Aug 12, 2000 this sequence version replaced gi:8810457.  
\* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 41526: contig of 41526 bp in length  
\* 41527 41726: gap of unknown length  
\* 41727 199551: contig of 157825 bp in length.

FEATURES  
source Location/Qualifiers  
1..199551  
/organism="Plasmodium falciparum"  
/db\_xref="taxon:5833"  
/chromosome="12"  
/clone="PFYACB8-95"  
/clone="3D7"

BASE COUNT 83197 a 17543 c 19759 g 78852 t 200 others  
ORIGIN

Query Match 8.4%; Score 52.6; DB 2; Length 199551;  
Best Local Similarity 49.1%; Pred. No. 0.21;  
Matches 139; Conservative 0; Mismatches 144; Indels 0; Gaps 0;  
QY 340 AAAATGAAATCATATAATACGAGGTAATTCGAGTTTAAATACACTAAATCTTCCAA 399  
Db 130728 AAAAGAAAAAAGATAAAAGAAAAAGAAAAAGAAAAAGAAATTAAGATATCCAA 130787  
QY 400 ATCAACCCCAATCTTCCACCAAGCTGGCAGATGTGGAAGTTATGGGAAAAAGGATAAAA 459  
Db 130788 ATGGACGTAAGGAAAAAGACAAGTTAGATGACCTAACTGAGGAAGATAAAATAAGAGCA 130847  
QY 460 TGGTTGATGGTGAAGCAAGTAATCACTTAAATAATGACATTATTCAGATATCTCGAAGAAA 519  
Db 130848 ATGGTTGTATGAAGAAGTGAAGAAAAAGAAATGAGAATTCGTTGATTTATTAAAAAGG 130907  
QY 520 TTAGAAACAAATGTAAAGTATCTGCATTGATAAATATGCGCTTAGCCATTTTCCAAATAT 579  
Db 130908 ATAGTAATACATCTGAACAACCTTAACAATAGCAATAATAGCTTGAATATTAGAAAAATA 130967  
QY 580 CTAATTGCTCACTCAAGTAAAAAAMAAAAAAMAAAAAAMAACTCGA 622  
Db 130968 ATAAAAATAGAAAAAATAAATAATAGAAAAATGAAAAACATGA 131010

RESULT 13  
DDISASIA 2145 bp mRNA linear INV 27-APR-1993  
LOCUS D.dictyostelium GTP-binding protein (SAS1) gene, complete cds.  
DEFINITION M34456  
ACCESSION M34456  
VERSION 1  
KEYWORDS GTP-binding protein.  
SOURCE D.dictyostelium, cDNA to mRNA.  
ORGANISM Dictyostelium discoideum  
REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.  
AUTHORS Saxe,S.A. and Kimmel,A.R.  
TITLE SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum with sequence similarities to essential genes in Saccharomyces cerevisiae  
JOURNAL Mol. Cell. Biol. 10, 2367-2378 (1990)  
MEDLINE 90220623  
FEATURES Location/Qualifiers  
1..2145

CDS

/organism="Dictyostelium discoideum"  
/db\_xref="taxon:44689"  
1095..1706  
/note="GTP-binding protein (SAS1)"  
/codon\_start=1  
/protein\_id="AAA33248.1"  
/db\_xref="GI:167875"  
/translation="MTSPATNKPAAAYDFLVKLLIGDSGVKSCLLLRFSQGSFTPSF  
IATIGIDFKIRTIIELEGKRIKIQIWDTAGQERFRTTAYYRGACGILLVYDVTDEKS  
FGSIRNIRNIEHSDSVNKMIGNKCDMPKVKVDSSRCKSLADEYGIKPLETSK  
NSVYVEAFIQLAKDIKKRMIDTPNDPHITICITPNKKNTCC"  
BASE COUNT 929 a 249 c 238 g 727 t 2 others  
ORIGIN

Query Match 8.4%; Score 52.2; DB 3; Length 2145;  
Best Local Similarity 47.6%; Pred. No. 0.6;  
Matches 150; Conservative 0; Mismatches 165; Indels 0; Gaps 0;  
QY 305 AACAACTCAAAGCCAGCTACATATCTTAACCAAGAAAAATGAAATCATATAATACAGGA 364  
Db 415 AAAATCTTAAAAACCAGAAAGAAATAATAATAATAATAATAATAATAATAATAATAA 474  
QY 365 AGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCAATCTTTCACGCGAAGCT 424  
Db 475 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 534  
QY 425 GGCAGATGTGAAAGTTATGGGAAAAAGGATAAAATGGTTGATGGTGAGCAAGTAATCAC 484  
Db 535 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 594  
QY 485 TAAATATGACATATTGAAGATACCTTCGAAAGAAAAATAGAAACAAATTAAGATATCTG 544  
Db 595 AAAAAAGAAATAGCTTTACATTTTAAATTAATAATAATAATAATAATAATAATAATAA 654  
QY 545 CATTGATAATATGGCCCTTAGCCATTTCCAAATATCTAAATTTGCTCACTCAAGTAAAAA 604  
Db 655 TATTGAAGTATATCAATAGTTTAAATTTAATTTATTGTTTATTATAAAAAAATAAAAA 714  
QY 605 AAAAAAAMAAAAACT 619  
Db 715 AAAAAAAMAAATATT 729

RESULT 14  
AF361475 5191 bp DNA linear INV 03-MAY-2001  
LOCUS Dictyostelium discoideum double histidine kinase Dhkd (dhkd) gene, complete cds.  
DEFINITION AF361475.1 GI:13937002  
ACCESSION AF361475  
VERSION 1  
KEYWORDS Dictyostelium discoideum.  
SOURCE Dictyostelium discoideum.  
ORGANISM Dictyostelium discoideum  
REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.  
AUTHORS 1 (bases 1 to 5191)  
TITLE Anjard,C. and Loomis,W.F.  
JOURNAL The histidine kinases of Dictyostelium  
(in) Inouye,M. and Dutta,R. (Eds.);  
HISTIDINE KINASES IN SIGNAL TRANSDUCTION;  
Academic press, San Diego (2001) In press  
REFERENCE 2 (bases 1 to 5191)  
AUTHORS Anjard,C. and Loomis,W.F.  
TITLE Direct Submission  
JOURNAL Submitted (15-MAR-2001) Biology 0368, UCSD, 9500 Gilman Drive, La Jolla, CA 92093, USA  
FEATURES Location/Qualifiers  
1..5191  
/organism="Dictyostelium discoideum"  
/db\_xref="taxon:44689"  
1..5191  
/gene="dhkd"  
join(<436..1082,1162..>5155)  
/gene="dhkd"



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:12:43 ; Search time 137.032 Seconds  
(without alignments)  
7805.766 Million cell updates/sec

Title: US-09-881-556-4

Perfect score: 623

Sequence: 1 cctcaatggtgaacatcaa.....aaaaaaaaaaaaactcgag 623

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Capext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
C 1	52	8.3	6183 24	ABL33197 Human immune syste
C 2	52	8.3	9539 22	AAS45347 Chemically pretrea
C 3	51.8	8.3	6915 24	ABL32406 Human immune syste
C 4	50.8	8.2	6494 24	ABL33393 Human immune syste
C 5	50.2	8.1	1757 24	ABL34421 Human immune syste
C 6	50.2	8.1	7306 24	ABL33637 Human immune syste
C 7	49.8	8.0	14006 24	ABL33958 Human immune syste
C 8	49.5	8.0	5493 24	AAS61386 Human gene regulat
C 9	49	7.9	5283 22	AAS46539 Tumour suppressor

C 10	48.8	7.8	6145	24	ABL32972	Human immune syste
C 11	48.6	7.8	10619	22	AAS45296	Chemically pretrea
C 12	48.6	7.8	10619	24	ABL32070	Human immune syste
C 13	48.4	7.8	3567	21	AAA70117	Plasmodium falcipa
C 14	48.4	7.8	5452	24	ABL33149	Human immune syste
C 15	48.4	7.8	6767	22	AAS46508	Tumour suppressor
C 16	48.2	7.7	6485	22	AAS46559	Tumour suppressor
C 17	48.2	7.7	6485	24	ABL33808	Human immune syste
C 18	48.2	7.7	7479	24	AAS63345	Chemically pretrea
C 19	48	7.7	7924	24	ABL34139	Human immune syste
C 20	47.8	7.7	8254	24	ABL32864	Human immune syste
C 21	47.8	7.7	13249	24	ABL32116	Human immune syste
C 22	47.6	7.6	819	22	AA194974	Human neuroblastom
C 23	47.6	7.6	8634	24	ABL33056	Human immune syste
C 24	47.2	7.6	7134	24	ABL32482	Human immune syste
C 25	47	7.5	6035	24	ABL33525	Human immune syste
C 26	47	7.5	6277	22	AAS46728	Tumour suppressor
C 27	47	7.5	6277	24	ABL34040	Human immune syste
C 28	46.8	7.5	5379	24	ABL33677	Human immune syste
C 29	46.8	7.5	5379	24	ABL34577	Human metastasis a
C 30	46.8	7.5	5413	22	AAS46694	Tumour suppressor
C 31	46.8	7.5	5454	21	AAA70236	Plasmodium falcipa
C 32	46.8	7.5	6070	24	ABL33679	Human immune syste
C 33	46.8	7.5	6070	24	ABL34579	Human metastasis a
C 34	46.8	7.5	6591	22	AAS46283	Tumour suppressor
C 35	46.8	7.5	7571	24	ABL32527	Human immune syste
C 36	46.8	7.5	9209	24	ABL34427	Human immune syste
C 37	46.8	7.5	9219	22	AAS46808	Tumour suppressor
C 38	46.8	7.5	14429	24	ABL34242	Human immune syste
C 39	46.8	7.5	17934	24	ABL33719	Human immune syste
C 40	46.6	7.5	5893	24	ABL32859	Human immune syste
C 41	46.6	7.5	6419	24	ABL32267	Human immune syste
C 42	46.4	7.4	4661	24	AAS61136	Human gene regulat
C 43	46.4	7.4	13606	22	AAS45457	Chemically pretrea
C 44	46.4	7.4	13606	22	AAS46561	Tumour suppressor
C 45	46.4	7.4	13606	24	ABL33810	Human immune syste

ALIGNMENTS

RESULT 1

ABL33197/C

ID ABL33197 standard; DNA: 6183 BP.

XX ABL33197;

XX ABL33197;

XX 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 1170.

Human; immune system disease; cytosine methylation; antiasthmatic; antiarteriosclerotic; antianaemic; cytostatic; nootropic; neuroprotective; anti-HIV; anticonvulsant; ophthalmological; antirheumatic; antiarthritic; antidiabetic; antipsoriatic; antineoplastic; cancer; eye disease; arteriosclerosis; anaemia; acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy; neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene; ds.

OS Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.





Db 4724 ATACCTAATAAAACCTCTAAATAACAAAACTTAAATTAATAATCTTTTATCTACAT 4665  
QY 357 AATCAGGAAGTAAATTCGAGTTTATATACACTAAATCT--TCCAAATCAACCAATCTTT 414  
Db 4664 ATTAATAATAATTAATTTATTTATTTCTTAACAAATTTTCTCTAAATTTCTCAAACTAT 4605  
QY 415 CAGGCAGCTGGCAGATGCGAAAGTTATGGGGAAGATAAATGGTTGATGGTGAGC 474  
Db 4604 CAATATATTAAAAAAAATTAATTTATTTCTACAACTTCTCTAAATTTATATATATAATAA 4545  
QY 475 AAGTATCATCAATAATGACATTAATGAGATCTTCGAAAGAAATAGAAACAAATCT 534  
Db 4544 ATAAATAATACCAATATAATATAATTAACAACTTAAAAAAAACAATACAATAAACA 4485  
QY 535 AAGTATCTGCTATGATAAATATGGCTTTAGCCATTTCCAAATATCTAAATTTGCAACTC 594  
Db 4484 AATATCTTATATCAAAATAAATAAATAATAAATATCTCTCACTAAATAATAAACA 4425  
QY 595 AAGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 622  
Db 4424 CAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4397

RESULT 5  
ABL34421/c  
ID ABL34421 standard; DNA; 1757 BP.  
XX AC ABL34421;  
XX DT 26-MAR-2002 (first entry)  
XX DE Human immune system associated gene SEQ ID NO: 2394.  
XX KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytostatic; neutropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX OS Homo sapiens.  
XX PN WO200200928-A2.  
XX PD 03-JAN-2002.  
XX PF 02-JUL-2001; 2001WO-EP07537.  
XX PR 30-JUN-2000; 2000DE-1032529.  
XX PR 01-SEP-2000; 2000DE-1043826.  
XX PA (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX DR WPI; 2002-130909/17.  
XX PT Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -  
XX PS Claim 1; SEQ ID NO 2394; 32pp + Sequence Listing; German.  
XX CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

CC diseases. The present sequence is a gene of the invention.  
XX SQ Sequence 1757 BP; 396 A; 83 C; 464 G; 813 T; 1 other;  
Query Match 8.1%; Score 50.2; DB 24; Length 1757;  
Best Local Similarity 60.7%; Pred. No. 0.027; 53; Indels 0; Gaps 0;  
Matches 82; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 483 ACTAAAAATGACATTTATTGAGATACCTTCGAAAGAAATTTAGAAACAAATGTAAAGTATC 542  
Db 1683 ATTAAAAATATTACATTTAAATTAATCCAAACAAAAATATAATAATATAAATACTAG 1624  
QY 543 TGCATTGATAAATATGGCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAAA 602  
Db 1623 ATAAATCTCAAAATTAATATACTAAAAAATATCTAAATAACAAATTTTACTATAAATAAAT 1564  
QY 603 AAAAAAATAAAAAA 617  
Db 1563 ATAAAAATAAAAAA 1549  
RESULT 6  
ABL33637/c  
ID ABL33637 standard; DNA; 7306 BP.  
XX AC ABL33637;  
XX DT 26-MAR-2002 (first entry)  
XX DE Human immune system associated gene SEQ ID NO: 1610.  
XX KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytostatic; neutropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX OS Homo sapiens.  
XX PN WO200200928-A2.  
XX PD 03-JAN-2002.  
XX PF 02-JUL-2001; 2001WO-EP07537.  
XX PR 30-JUN-2000; 2000DE-1032529.  
XX PR 01-SEP-2000; 2000DE-1043826.  
XX PA (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX DR WPI; 2002-130909/17.  
XX PT Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -  
XX PS Claim 1; SEQ ID NO 1610; 32pp + Sequence Listing; German.  
XX CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.  
XX SQ Sequence 7306 BP; 2019 A; 93 C; 1645 G; 3549 T; 0 other;

Query Match 8.1%; Score 50.2; DB 24; Length 7306;  
Best Local Similarity 47.1%; Pred. No. 0.037;  
Matches 154; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 291 AATCACTTCTACAGACAACTCAAGCCAGCTACATATCTTCAACCAAGAAATGAAAT 350  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 6509 AATCCCACTACTCGAATAAATAATAATAATAATAATAATAATAATAATAATA 6450  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 351 CATATAAATCAGGAAGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCAAT 410  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 6449 TTCAATAAATAAATATATATTCACACTACACTAACTTAATAACAACCACTTTATC 6390  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 411 CTTTCAGCAGCTGGCAGATGTGGAAAGTTATGGGAAAGGATAAAATGGTTGATGGT 470  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 6389 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6330  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 471 GAGCAAGTATCACTAAATAATGACATTTATGAGATATCTCGAAGAAATAGAAACAAA 530  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 6329 AAA 6270  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 531 ATCTAAAGTATCTGCAATGATAATATGGCCTTAGCCATTTCCAAATATATCTAAATGTCA 590  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 6269 AAA 6210  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 591 ACTCAAGTAAAAAAAAAAAAAAAAAAAAA 617  
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 6209 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6183  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7  
ABL33958/c  
ID ABL33958 standard; DNA; 14006 BP.  
XX ABL33958;  
XX  
XX  
XX 26-MAR-2002 (first entry)  
DE Human immune system associated gene SEQ ID NO: 1931.  
XX  
XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytosine; cytosine; cytosine;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriasis;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200200928-A2.  
XX  
XX 03-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
XX  
XX 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIC-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
PI WPI; 2002-130909/17.  
XX  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -  
XX  
XX Claim 1; SEQ ID NO 1931; 32pp + Sequence Listing; German.  
XX  
XX The present invention provides a number of human immune system associated

CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.  
XX

SQ Sequence 14006 BP; 3278 A; 155 C; 3257 G; 7313 T; 3 other;  
Query Match 8.0%; Score 49.8; DB 24; Length 14006;  
Best Local Similarity 47.3%; Pred. No. 0.053;  
Matches 150; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTCAACCAAGAAATGAAATCATATAATC 360  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2441 AA 2382  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 361 AGAAGGTAAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCAATCTTTCAGCA 420  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 2381 ACCAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2322  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 421 AGCTGGCAGATGTGGAAAGTTATGGGAAAGGATAAAATGGTGGTGAAGCACTAA 480  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 2321 AAAAAAAAAAAAAAAAAACGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2262  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 481 TCACTAAATATGACATTTATTTGAAGATATCTCGAAAGAAATTAGAAACAAATGTAAAGTA 540  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2261 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2202  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 541 TCTGCATTGATTAATATATGGCCTTAGCCATTTCCAAATATATCTTAATTTCTCACTCACTAA 600  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 2201 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2142  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 601 AAAAAAAAAAAAAAAAAAAAA 617  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2141 AAAAAAAAAAAAAAAAAAAAA 2125  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8  
AAS61386/c  
ID AAS61386 standard; DNA; 5493 BP.  
XX  
XX AAS61386;  
XX  
XX 29-JAN-2002 (first entry)  
DE Human gene regulation-associated gene oligonucleotide #341.  
XX  
XX Human; Gene regulation-associated gene; severe combined immunodeficiency;  
KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;  
KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;  
KW renal disease; preclampsia; cardiac allograft vascular disease;  
KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;  
KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;  
KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.  
XX  
XX Homo sapiens.  
XX  
XX WO200177375-A2.  
XX  
XX 18-OCT-2001.  
XX  
XX 06-APR-2001; 2001WO-EP03968.  
XX  
XX  
XX 06-APR-2000; 2000DE-1019058.  
XX  
XX 07-APR-2000; 2000DE-1019173.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
XX  
XX 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIC-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;

XX WPI: 2002-017470/02.  
XX  
PT New nucleic acid sequences from chemically modified genes associated  
PT with gene regulation, useful for analysing cytosine methylations for  
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency  
PT disease -  
XX  
PS Disclosure: SEQ ID No 349; 26pp; English.  
XX  
XX The invention relates to 224 nucleic acid sequences comprising at least  
CC 18 bases of a chemically pretreated gene associated with gene regulation  
CC selected from 43 known genes (or complementary sequences). The  
CC chemical pretreatment converts cytosine bases unmethylated at the  
CC 5-position to uracil or another base with hybridisation behaviour  
CC dissimilar to cytosine, to enable analysis of cytosine methylations.  
CC The DNA sequences, oligomers (or sets/arrays) and method are  
CC useful in the diagnosis of diseases (or predisposition to diseases)  
CC associated with gene regulation and in therapy of such diseases, by  
CC enabling analysis of the cytosine methylation patterns of such genes,  
CC kits are provided. They are especially useful in diagnosis  
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac  
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,  
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,  
CC precociousness, graft versus-host disease. The present sequence is a  
CC sequence included in the sequence data for this specification and is  
CC associated with the human gene regulation-associated genes.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 5493 BP; 915 A; 149 C; 1631 G; 2798 T; 0 other;  
Query Match 8.0%; Score 49.6; DB 24; Length 5493;  
Best Local Similarity 47.4%; Pred. No. 0.048;  
Matches 148; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
QY 306 ACAACTCAAGCCAGTCATATCTTACCGAAGAAATGAAATCATATAATCAGGAA 365  
DB 4424 ACNAAAACAAAATAATACACTTAAAAAATAAAAAACAAAAAATAAAAAA 4365  
QY 366 GGTAAATTCGAGTTTAAATACACTAATCTTCCAAATCAACCCAAATCTTCACGCAAGCTG 425  
DB 4364 AAAATATATATATATACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 4305  
QY 426 GCAGATGTGGAAGTTATGGGAAAGGATATAAATGTTGATGGTGCAAGTAACTACT 485  
DB 4304 AAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAA 4245  
QY 486 AAAATGACATTATGAAGATCTTTCGAAAGAAATTAGAAACAAATGTAAGTATCTGC 545  
DB 4244 ACAATAAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACA 4185  
QY 546 ATTGATATAATGGCTTATAGCCATTTCAATATCTAATTTGTCACATCAAGTAAAAAA 605  
DB 4184 ATTAATAAACAAACCAACAAAAAATAAAAAAATAAAAAAACAACAAAAAACAATA 4125  
QY 606 AAAAAAATAAAAA 617  
DB 4124 AAAAAAATAAAAA 4113  
RESULT 9  
ID AAS46539/c  
XX AAS46539 standard; DNA; 5283 BP.  
XX  
AC AAS46539;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Tumour suppressor gene derived chemically modified sequence #261.  
XX

KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;  
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
KW cytosine methylation; ds.  
OS Homo sapiens.  
XX WO200168912-A2.  
PN  
XX 20-SEP-2001.  
PD  
XX  
XX 15-MAR-2001; 2001WO-EP02955.  
PF  
XX 15-MAR-2000; 2000DE-1013847.  
PR  
XX 06-APR-2000; 2000DE-1019058.  
PR  
XX 30-JUN-2000; 2000DE-1019173.  
PR  
XX 01-SEP-2000; 2000DE-1032529.  
PR  
XX (EPIG-) EPIGENOMICS AG.  
PA  
XX Olek A, Piepenbrock C, Berlin K;  
PI  
XX WPI: 2001-602752/68.  
DR  
XX Fragments of chemically modified genes associated with tumour suppressor  
PT genes and oncogenes, useful in designing primers and probes for  
PT analysing diseases associated with cytosine methylation state e.g.  
PT cancer -  
XX  
PS Claim 1; SEQ ID No 261; 27pp; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence of 18  
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
CC bisulphite, of genes associated with tumour suppression and  
CC oncogenes having a sequence taken from 536 (actually 533 since  
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences  
CC (SS) and sequences complementary to (SS). The nucleic acid may be a  
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may  
CC form part of a set of probes for detecting the cytosine methylation state  
CC and/or single nucleotide polymorphisms and also to be used in an  
CC array for analysing diseases associated with CpG dinucleotides e.g.  
CC cancers and tumours. The probes can also be used in a method for  
CC ascertaining genetic and/or epigenetic parameters for the diagnosis  
CC and/or therapy of existing diseases or the predisposition to specific  
CC diseases, by analysing cytosine methylations. The parameters may be  
CC compared to another set of genetic and/or epigenetic parameters, the  
CC differences serving as basis for diagnosis and/or prognosis events which  
CC are disadvantageous to patients. The present sequence is one of the  
CC 533 genomic sequences derived from tumour suppressor genes and  
CC oncogenes.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 5283 BP; 1113 A; 151 C; 1299 G; 2720 T; 0 other;  
Query Match 7.9%; Score 49; DB 22; Length 5283;  
Best Local Similarity 47.3%; Pred. No. 0.066;  
Matches 148; Conservative 0; Mismatches 165; Indels 0; Gaps 0;  
QY 305 AACAACTCAAGCCAGCTACATATCTTCAACAGAAATGAAATCATATAATCAGGA 364  
DB 2358 AAATAAAAAAACAACAAATCAAAAAATATCCAAATATTTCCATTCATATAAACTCCA 2299  
QY 365 AGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCAAATCTTTCACGCAAGCT 424  
DB 2298 AAACAAACAAACCAACCACTATACAAACAAATCAAAATTAATTAATCTTAAACCTAAAAA 2239  
QY 425 GGCAGATGTGAAAGTTATGGGAAAGGATATAAATGTTGATGGTGCAAGTAATCAC 484  
DB 2238 AAAATAAAATTAACATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAC 2179

QY 485 TAAAAATGACATTATTGAGTACTTCGAAGAAGTAATAGAACCAAAATGTAAAGTATCTG 544  
 Db 2178 CATCAATCCACCACCGATACACTCAACCTAAATAACAAAAAATCTATCTCAAA 2119  
 QY 545 CATTTGATAAATATGGCCTTAGCCATTTCGCAATATCTCAAAATGTCGCAACTCAAGTAAAAAA 604  
 Db 2118 AAAAAAATAATATCAATATATTAATCACTATATCACTATATCCACACTTTAAAAAAC 2059  
 QY 605 AAAAAAATAA 617  
 Db 2058 CAAAAAATAA 2046  
 RESULT 10  
 ABL32972/c  
 ID ABL32972 standard; DNA; 6145 BP.  
 XX  
 AC ABL32972;  
 XX  
 DT 26-MAR-2002 (first entry)  
 DE  
 XX Human immune system associated gene SEQ ID NO: 945.  
 XX Human; Immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; antihaemic; cytostatic; nootropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;  
 KW antithumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 KW gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200928-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 02-JUL-2001; 2001WO-EP07537.  
 XX  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2002-130909/17.  
 XX  
 PT Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation -  
 XX  
 XX Claim 1; SEQ ID NO 945; 32pp + Sequence Listing; German.  
 XX  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 XX  
 SQ Sequence 6145 BP; 1493 A; 99 C; 1309 G; 3244 T; 0 other;

Query Match 7.88; Score 48.8; DB 24; Length 6145;  
 Best Local Similarity 47.28; Pred. No. 0.076;  
 Matches 149; Conservative 0; Mismatches 167; Indels 0; Gaps 0;  
 QY 301 ACAGACAAACTCAAGCCAGCTACATATCTCAACCAAGAAATGAAAAATCAATAATC 360  
 Db 5661 AATTAATATATACAAAAAACAATAATTAATAAAAAAATAAAAAAATA 5802

QY 361 AGGAAGTAAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCAATCTTTTCAGCGA 420  
 Db 5801 AATAACTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5742  
 QY 421 AGCTGGCAGATGTGGAAGTTATCGGGAAGAAGGATNAATGGTTGATGGTGAGCAGTAA 480  
 Db 5741 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5682  
 QY 481 TCACATAAAATGACATTATTCAAGATATCTTCGAAAGAAATTTAGAAACAAATGTAAAGTA 540  
 Db 5681 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5622  
 QY 541 TCTGCATTGATAATATGCGCTTAGCCATTTCAGCCATTTCCAATATCTTAATTTGTCMACTCAAGTAA 600  
 Db 5621 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5562  
 QY 601 AAAAAAATAA 616  
 Db 5561 AATAAATAAATAAATAA 5546  
 RESULT 11  
 AAS45296/c  
 ID AAS45296 standard; DNA; 10619 BP.  
 XX  
 AC AAS45296;  
 XX  
 DT 18-DEC-2001 (first entry)  
 DE  
 XX Chemically pretreated genomic DNA associated with cell cycle #1.  
 XX Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;  
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;  
 KW graft-versus-host disease; glomerular disease; lowy body disease; cancer;  
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;  
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;  
 KW PCR primer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200168911-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 15-MAR-2001; 2001WO-EP02945.  
 XX  
 PR 15-MAR-2000; 2000DE-1013847.  
 PR 06-APR-2000; 2000DE-1019058.  
 PR 07-APR-2000; 2000DE-1019173.  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2001-602751/68.  
 XX  
 PT Designing primers and probes for analysing diseases associated with  
 PT cytosine methylation state e.g. arthritis, cancer, aging,  
 PT arteriosclerosis comprising fragments of chemically modified genes  
 PT associated with cell cycle -  
 XX  
 XX Claim 1; SEQ ID NO 1; 28pp; English.  
 XX  
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA  
 CC molecules associated with the cell cycle and specific PCR primers of the  
 CC invention. The sequences are useful for detecting the methylation state  
 CC of all CpG dinucleotides in a sequence and therefore for analysing  
 CC associated diseases. By analysing cytosine methylations in the pretreated  
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy  
 CC of existing diseases or the predisposition to specific diseases can be







PT Plasmodium falciparum, useful as antimalarial vaccines and in the  
PT diagnosis of P.falciparum infection -  
XX  
PS Disclosure: Page 470-471; 577pp; English.  
XX  
CC The present invention describes proteins and their fragments (I) encoded  
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.  
CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)  
CC vaccines against P. falciparum infection comprising (I) or (II).  
CC (I) and (II) are useful for the development of vaccines against  
CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal  
CC antibody raised to immunogens comprising the sequences of (I), are  
CC useful in the detection of infection with P. falciparum. Furthermore,  
CC (I) especially when they are rifins or secreted or membrane proteins)  
CC can aid the identification of drugs to treat or prevent P. falciparum  
CC infection, or they can be used to identify drug resistance in  
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the  
CC subsequent identification of proteins encoded by it will help to expand  
CC our understanding of parasite biology, a process hampered by the  
CC complexity of the parasitic lifecycle, and provide new targets for  
CC vaccine and drug development. Parasite resistance to drugs and mosquito  
CC resistance to insecticides have led to a resurgence of malaria in many  
CC parts of the world, and there is a pressing need for vaccines and new  
CC drugs. AAA70078 and AAB18144 to AAB18352 represent nucleotide  
CC and protein sequences given in the present invention, but which are not  
CC specifically mentioned within the specification.  
XX  
SQ Sequence 3567 BP; 1641 A; 269 C; 473 G; 1184 T; 0 other;  
Query Match 7.8%; Score 48.4; DB 21; Length 3567;  
Best Local Similarity 47.9%; Pred. No. 0.083;  
Matches 139; Conservative 0; Mismatches 151; Indels 0; Gaps 0;  
Qy 328 ATCTTAACCAAGAAATGAAATCATATATATCAGGAGGTAATTCGAGTTTAAATACAC 387  
Db 1373 ATATTAATTAATGATGATATTAATTAATTAATGATGATTAATTAATTAATG 1432  
Qy 388 TAAATCTTCCAAATCAACCCCAATCTTTCACCCAGCTGGCAGAGTGGAAAGTTATGGG 447  
Db 1433 ATATATTAATTAATTAATGATGATTAATTAATTAATGATGATTAATTAATTAAT 1492  
Qy 448 AAAAGGATAAAATGGTGTGCTGAGCAAGTAATCACTTAAATTAATGACATTAATGAAGATA 507  
Db 1493 ATGATATATTAATTAATTAATGATGATTAATTAATTAATGATGATTAATTAATTA 1552  
Qy 508 CTTCGAAGAAATAGAAACAAATGTAAAGTATCTGCATTGATAAATATGCGCTTAGCC 567  
Db 1553 ATAATGATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1612  
Qy 568 ATTTCCAAATATCTAAATTTGCACTCAAGTAAAAAATTAATTAATTAATTAATTAAT 617  
Db 1613 AGATCCTTTTAAACACACATCCCAATGTACAGATGTTTAAAGAA 1662  
RESULT 14  
ABL33149/G  
ID ABL33149 standard; DNA: 5452 BP.  
XX  
XX ABL33149;  
AC  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human immune system associated gene SEQ ID NO: 1122.  
XX  
KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianemic; cytosine; cytosine; cytosine;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriasis;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.

OS Homo sapiens.  
XX  
PN WO200200928-A2.  
XX  
PD 03-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
XX  
PR 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
PA (EPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-130909/17.  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -  
XX  
PS Claim 1; SEQ ID NO 1122; 32pp + Sequence Listing; German.  
XX  
CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.  
XX  
SQ Sequence 5452 BP; 1328 A; 55 C; 1065 G; 3004 T; 0 other;  
Query Match 7.8%; Score 48.4; DB 24; Length 5452;  
Best Local Similarity 47.9%; Pred. No. 0.091;  
Matches 139; Conservative 0; Mismatches 151; Indels 0; Gaps 0;  
Qy 327 TATCTTAACCAAGAAATGAAATCATATTAATCAGGAGGTAATTCGAGTTTAAATACA 386  
Db 5201 TATCTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 5142  
Qy 387 CTAAATCTTCCAAATCAACCCCAATCTTTCACGCAAGCTGGCAGATGCGAAAGTTATGGG 446  
Db 5141 AAAAACAATACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5082  
Qy 447 GAAAGGATAAAATGGTGTGAGCAAGTAATCACTTAAATTAATGACATTAATGAAGAT 506  
Db 5081 AACAAATATCATCTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5022  
Qy 507 ACTTCGAAAGAAATAGAAACAAATGTAAAGTATCTGCATTGATAAATATGCGCTTAGC 566  
Db 5021 CAATAACACACTTAATCAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 4962  
Qy 567 CATTTCCAAATATCTAAATTTGCACTCAAGTAAAAAATTAATTAATTAATTAATTAAT 616  
Db 4961 AATTATCCAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 4912  
RESULT 15  
AAS46608/G  
ID AAS46608 standard; DNA: 6767 BP.  
XX  
XX AAS46608;  
AC  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Tumour suppressor gene derived chemically modified sequence #330.  
XX  
KW Human; tumour suppressor gene; oncogene; antitumour; cytosine;  
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
KW cytosine methylation; ds.

OS Homo sapiens.

XX PN WO200168912-A2.

XX  
PD  
20-SEP-2001.

15-MAR-2001

15 JUL 2001, 2001NO 2502300  
XX  
15 MAR 2000, 2000M 1033013  
DB

PR 15-MAR-2000; 2000DE-1013847.  
PR 06-APR-2000; 2000DE-1019058.

PR 07-APR-2000; 2000DE-1019173.  
PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.  
XX

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C,

XX  
DR WPI; 2001-602752/68.

XX  
PT  
Fragments of chemical

PT analysing diseases associated with cytosine methylation state e.g. cancer -

PS Claim 1; SEQ ID No 330; 27pp; English.

The invention relates to a nucleic acid comprising a sequence of 18 bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and 500 are missing from the sequence listing) sequences (Ss) and sequences complementary to (Ss). The nucleic acid may be a peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms and also to be used in an array for analysing diseases associated with CpG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases, by analysing cytosine methylations. The parameters may be compared to another set of genetic and/or epigenetic parameters, the differences serving as basis for diagnosis and/or prognosis events which are disadvantageous to patients. The present sequence is one of the 533 genomic sequences derived from tumour suppressor genes and oncogenes. Sequences with even numbered Seq ID numbers are the complementary sequence of the corresponding odd numbered sequence (e.g. ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence is missing).

CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

xx  
SQ Sequence 6767 BP; 1446 A; 207 C; 1601 G; 3513 T; 0 other;

Query Match 7.8%; Score 48.4; DB 22; Length 6767;  
Best Local Similarity 46.7%; Pred. No. 0.095;  
Matches 154; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 288 ACCAATCACCTTCTACAGAACAACACTCAAAGCCAGCTACATATCTCTAACCAAGAAATGAA 347

Db 2780 AACAAACAACATTAATTAATAAACCACAAAAACAAAAACACAAAAACAT 2721

QY 348 AATCATAATAATCAGSAAGGTAATTTCGAGTTTAAATACACTAAATCTTTCCAAATCAACCC 407

Db 2720 AAAAAAATAACAAAACCAAAACAAAAATCAAAAACAAAAACAAAAAAACCAACA 2661

QY 408 AATCTTTCACGCAAGCTGGCAGATGTGGAAGTTATGGCGAAAGATAAAATCGTTGAT 467

Db 2660 CAAACACAAAAAATCACAACCAAAACAAACGACAAAAATACAAAAACAAAAA 2601

QY 468 GGTGAGCAACTAATCACTAAAAATGACATTATTGAAGATCTTCGAAAGAAATTAGAAAC 527

•

Db	2500	AAAAAAAAAATACAAACCCAAAAACAATCAAAACCAAAATAAAAACAAAAAATAC	2541
Qy	528	AAATGTAAAGTATCTGCATTTGATAAATATGGCTTAGCCATTCCAAATATCTAAATTG	587
Db	2540	AAACCAAAAAACAATACAAAAACAATAACGAACCAAAATAAAAAANAATCA	2481
Qy	588	TCACTCTAGTAAAAAAAAAAAAAAAAAAAA	617
Db	2480	AAACCAAAAAAATAATATAAAACAAAAATA	2451

QY 528 AAAATGTAAGTATCTGCATTGATATAATATGCGCTTAGCCATTCCAAAATATCTAAATTG 587

Db 2540 AAAACCAAAACAAATACAAAACAAAAATACGAACCAAAAAATAAAAAAATCA 2481

Qy 588 TCAACTCAAGTAAAAAAAAAAAA 617

Db 2480 AAAACAAAAAAATATATAACAAAAATA 2451

Search completed: November 1, 2002, 23:24:25  
Job time : 218.032 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:17:03 ; Search time 31.0794 Seconds  
(without alignments)  
4923.835 Million cell updates/sec

Title: us-09-881-556-4  
Perfect score: 623  
Sequence: 1 cctcaatggtgaactcaa.....aaaaaaaaaaaaaactcgag 623

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues  
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents.NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	46.8	7.5	7218	1	US-08-232-463-14
C 2	43.2	6.9	5852	1	US-07-867-106-2
C 3	42.8	6.9	289	4	US-09-007-005-17
C 4	42.8	6.9	289	4	US-09-244-796-17
C 5	42.4	6.8	132331	3	US-09-128-155-16
C 6	41.2	6.6	19124	2	US-08-487-826B-13
C 7	40.4	6.5	2447	2	US-09-014-969-14
C 8	39.8	6.4	2619	5	PCT-US96-10521-17
C 9	39.8	6.4	2887	5	PCT-US96-10521-14
C 10	39.4	6.3	2649	2	US-08-718-864-1
C 11	39.4	6.3	2649	2	US-09-059-964A-1
C 12	39.4	6.3	2649	2	US-08-842-341-1
C 13	37.8	6.1	1886	4	US-08-936-165A-224
C 14	37.6	6.0	2019	3	US-09-040-843-3
C 15	37.6	6.0	2019	4	US-09-621-855-3
C 16	37.6	6.0	2599	3	US-09-040-843-1
C 17	37.6	6.0	2599	4	US-09-621-855-1
C 18	37.2	6.0	658	4	US-08-998-416-595
C 19	37	5.9	566	1	US-08-211-942-10
C 20	37	5.9	2836	3	US-08-747-221B-24
C 21	37	5.9	2836	3	US-08-747-221B-26
C 22	37	5.9	2836	4	US-09-005-051-24
C 23	37	5.9	2836	4	US-09-005-051-26
C 24	36.8	5.9	240	1	US-08-628-417-6
C 25	36.6	5.9	9636	4	US-08-323-170B-1
C 26	36.6	5.9	9636	4	US-08-954-441-1
C 27	36.2	5.8	2230	4	US-08-378-313-24

28	36.2	5.8	2672	3	US-09-214-564A-5	Sequence 5, Appli
29	36.2	5.8	2815	3	US-09-214-564A-1	Sequence 1, Appli
30	36	5.8	440	4	US-08-936-165A-24	Sequence 24, Appl
31	36	5.8	1493	1	US-08-340-820-24	Sequence 24, Appl
32	36	5.8	1493	1	US-08-593-535-24	Sequence 24, Appl
33	36	5.8	2445	6	5215909-9	Patent No. 5215909
34	35.8	5.7	2223	1	US-08-257-073-4	Sequence 4, Appli
35	35.6	5.7	1119	4	US-09-178-973B-7	Sequence 7, Appli
36	35.6	5.7	1119	4	US-09-354-243B-7	Sequence 7, Appli
37	35.4	5.7	1117	4	US-09-247-373B-33	Sequence 33, Appl
38	35.4	5.7	5852	1	US-07-867-106-2	Sequence 2, Appli
39	35.2	5.7	8920	2	US-08-446-855A-1	Sequence 1, Appli
40	35.2	5.7	8920	4	US-09-150-741-1	Sequence 1, Appli
41	35	5.6	473	1	US-08-764-100-16	Sequence 16, Appl
42	35	5.6	1066	1	US-08-157-101A-4	Sequence 4, Appli
43	35	5.6	1279	4	US-09-277-716-31	Sequence 31, Appl
44	35	5.6	4970	1	US-08-764-100-14	Sequence 14, Appl
C 45	35	5.6	4970	1	US-08-764-100-20	Sequence 20, Appl

## ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232.463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935.313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fis  
US-08-232-463-14

Query Match 7.5%; Score 46.8; DB 1; Length 7218;

Best Local Similarity 0.8%; Pred. No. 0.0057;  
Matches 3; Conservative 215; Mismatches 142; Indels 0; Gaps 0;

QY 262 AGAAGAGGATCGACTCAGACGTTTACCAATCCTTACAGAACAACTCAAGCCAG 321  
Db 1422 RRR 1363  
QY 322 CTACATATCTTACCAAGAAATGAAATCATATAATCAGGAAGTAATTCGAGTTTA 381  
Db 1362 RRR 1303  
QY 382 ATACATAATCTTCAATCAACCAATCTTTCAGCGAAGCTGGCAGATGGGAAGTT 441  
Db 1302 RRR 1243  
QY 442 ATGGGAAAAGGATAAATGGTGTGAGCACTCACTAAATGACATTTATG 501  
Db 1242 RRR 1183  
QY 502 AGATACTTCGAAAGAAATAGAAACAAATGTAAAGTATCTGCAATGATAATATGGCC 561  
Db 1182 RRR 1123  
QY 562 TTAGCAATTCCAATATCTAATCTCACTCACTCACTCACTCACTCACTCACTCG 621  
Db 1122 RRR 1063

RESULT 2  
US-07-867-106-2/c  
; Sequence 2, Application US/07867106  
; Patent No. 5389526  
; GENERAL INFORMATION:  
; APPLICANT: Slade, Martin B  
; APPLICANT: Chang, Andy C M  
; APPLICANT: Williams, Keith L  
; TITLE OF INVENTION: Improved plasmid Vectors for Cellular  
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526rls  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/867,106  
; FILING DATE: 19920625  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PJ 7187  
; APPLICATION NUMBER: PCT/AU90/00530  
; FILING DATE: 02-NOV-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feeney, Joanne Longo  
; REGISTRATION NUMBER: 35,134  
; REFERENCE/DOCKET NUMBER: RICE-0002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5852 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ANTI-SENSE: NO

FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2378..5038  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2378..5038  
US-07-867-106-2

Query Match 6.9%; Score 43.2; DB 1; Length 5852;  
Best Local Similarity 47.6%; Pred. No. 0.044; Indels 2; Gaps 1;  
Matches 159; Conservative 0; Mismatches 173;

QY 286 TTACCAATCCTTACAGAACAACTCAAGCCAGCTACATATCTTACCAAGAA--AA 343  
Db 2317 TTTTACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2258  
QY 344 TGAATAATCATATAATCAGGAAGGTAATTCGAGTTTAAATACATAAATCTTCCAATCA 403  
Db 2257 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2198  
QY 404 ACCCAATCTTCCACCAAGCTGCAGATCTGCAAAAGTTATGCGAAAGGATAAATGCT 463  
Db 2197 TAAATATAATTAATAATGTCATGCCAAACTGATAAATATTTGATATATATTCCTCAATAT 2138  
QY 464 TGATGGTGCAGCAAGTAATCAGTAATAAATGACATTTATGGAAGATACCTCGAAAGAAATAG 523  
Db 2137 TATAAATAAGGTATAATTAGATAGAGATAAATAAATAAATAAATAAATAAATAA 2078  
QY 524 AAACAAATGTAAGTAATCTGCAATGATAAATATGCGCTTAGCCATTTCCAATATCTAA 583  
Db 2077 AAAAAACCAAGTAATTAATATATTAATGAGGTTTTTTTTTTTTTTTTTTTTTTTT 2018  
QY 584 ATTGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 617  
Db 2017 TTTTTCATAAGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1984

RESULT 3  
US-09-007-005-17  
; Sequence 17, Application US/09007005B  
; Patent No. 6258558  
; GENERAL INFORMATION:  
; APPLICANT: Szostak, Jack W.  
; APPLICANT: Roberts, Richard W.  
; APPLICANT: Liu, Rihe  
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN  
; TITLE OF INVENTION: FUSIONS  
; FILE REFERENCE: 00786/350003  
; CURRENT APPLICATION NUMBER: US/09/007,005B  
; CURRENT FILING DATE: 1998-01-14  
; EARLIER APPLICATION NUMBER: 60/035,963  
; EARLIER FILING DATE: 1997-01-27  
; EARLIER APPLICATION NUMBER: 60/064,491  
; EARLIER FILING DATE: 1997-11-06  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 289  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Translation template  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(289)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-007-005-17

Query Match 6.9%; Score 42.8; DB 4; Length 289;  
Best Local Similarity 12.9%; Pred. No. 0.021;  
Matches 36; Conservative 98; Mismatches 146; Indels 0; Gaps 0;

QY 338 AGAAATGAAAAATCATAATAATCAGGAAGTAATTCGAGTTTTTATACACATAAATCTTCC 397

Db	3	RGRRCARARURURACRURURURACRARURURACRARURURGRNRNR	62
Qy	398	AAATCAACCCAAATCTTTTCACGCAAGCTGCAGATGTGGAAAGTTATGGGAAAAAGATAA	457
Db	63	RNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRS	122
Qy	458	AATGGTGTGATGGTGACAAAGTATCACTAAAAATGACATTTATGAAGATACATTCGAAAGA	517
Db	123	RNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRS	182
Qy	518	AATTAGAAACAAATGTAAAGTATCTGCATTGTATAATATGGCCTTAGCCATTTCCAAT	577
Db	183	RNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRSNRNRS	242
Qy	578	ATCTAAATTGCTCAACTCAAGTAAAAAAAAAAAAAAAAAAAAA	617
Db	243	RURCRURURGRCRCRCRUAAAAAAAAAAAAAAAAAAAAAAAAA	282

```

RESULT 4
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: FUSIONS
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

```

[illegible]

```

Qy   578 ATCTAAATTTGTCACACTCAAGTAAAAAAAAAAAAAAAAAAAAA 611
      : : : : | : |||||
Db   243 RURCRURRGRRCRGCRUAAAAAAAAAAAAAAAAAAAAA 282

RESULT 5
US-09-128-155-16
; Sequence 16, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 REL
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; EARLIER FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

```

[illegible]

RESULT 6  
 US-08-487-826B-13  
 : Sequence 13, Application US/08487826B  
 : Patent No. 5993827  
 : GENERAL INFORMATION:  
 : APPLICANT: Sim, Kim L.  
 : APPLICANT: Chitnis, Chetan  
 : APPLICANT: Miller, Louis H.  
 : APPLICANT: Peterson, David S.  
 : APPLICANT: Su, Xin-zhaun  
 : APPLICANT: Wellens, Thomas E.  
 : TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
 : AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
 : NUMBER OF SEQUENCES: 45  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Knobbe Martens Olson & Bear  
 : STREET: 620 Newport Center Drive 16th Floor  
 : CITY: Newport Beach  
 : STATE: California  
 : COUNTRY: US



```

; PRIOR APPLICATION DATA: IL 114,615
; APPLICATION NUMBER: IL 114,615
; FILING DATE: 16-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,986
; FILING DATE: 17-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 115,319
; FILING DATE: 14-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 116,588
; FILING DATE: 27-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117,932
; FILING DATE: 16-APR-1996
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; PCT-US96-10521-17

Query Match 6.4%; Score 39.8; DB 5; Length 2619;
Best Local Similarity 57.7%; Pred. No. 0.25;
Matches 71; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 495 ATTATTCAAGATACCTTCGAAGAATAATTAGAACAAATGTAAAGTATCTGCATTGATATAA 554
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 ATTATTTTAAATCATTAAGGAATTAAGTTATCTTTTAAATTTTAAAGTATCTTTTCCAAA 2549
QY 555 TATGGCCCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAAAAAATAAAAAA 614
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2550 CATTTTATAGATATAATTAATTTGATCTTTAAAAAATAAAAAAATAAAAAA 614
QY 615 AAA 617
|||
Db 2610 AAA 2612

RESULT 9
PCT-US96-10521-14
; Sequence 14, Application PC/TUS9610521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
; TITLE OF INVENTION: AND OTHER PROTEINS
; NUMBER OF SEQUENCES: 34
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10521
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,615
; FILING DATE: 16-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,986
; FILING DATE: 17-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 115,319
; FILING DATE: 14-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 116,588
; FILING DATE: 27-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117,932
; FILING DATE: 16-APR-1996

```

QY 495 ATTATTGAAGATACCTCGAAGAAATAGAACAAATGTAAAGATATCTGCATTGATATA 554  
DB 2529 ATTATTTAAATCATAGGAATTAAGTTATCTTTAAATTTAAAGTATCTTTTTCAAAAA 2588  
QY 555 TATGGCCTTAGCCATTCCAAATATCTAAATGTCAACTCAAGTAAACAAAAA 614  
DB 2589 CATTTTAAATAGATAAATAATATTTGATCTTAAAAA 2648  
QY 615 A 615  
DB 2649 A 2649

## RESULT 11

US-09-059-964A-1  
; Sequence 1, Application US/09059964A  
; Patent No. 5952228

## GENERAL INFORMATION:

; APPLICANT: van der Bruggen, Pierre; Mandruzzato, Susanna;  
; APPLICANT: Boon-Falleur, Thierry  
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding A Functional  
; TITLE OF INVENTION: Protein, The Protein So Encoded And Uses Thereof  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski, L.L.P..  
; STREET: 555 Fifth Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/059,964A  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/718,964  
; FILING DATE: 26-September-1996  
; APPLICATION NUMBER: 08/669,590  
; FILING DATE: 24-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5952228man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5442.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2649 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

## US-09-059-964A-1

Query Match 6.3%; Score 39.4; DB 2; Length 2649;  
Best Local Similarity 57.9%; Pred. No. 0.31;  
Matches 70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 495 ATTATTGAAGATACCTCGAAGAAATAGAACAAATGTAAAGATATCTGCATTGATATA 554  
DB 2529 ATTATTTAAATCATAGGAATTAAGTTATCTTTAAATTTAAAGTATCTTTTTCAAAAA 2588  
QY 555 TATGGCCTTAGCCATTCCAAATATCTAAATGTCAACTCAAGTAAACAAAAA 614  
DB 2589 CATTTTAAATAGATAAATAATATTTGATCTTAAAAA 2648  
QY 615 A 615  
DB 2649 A 2649

## RESULT 13

US-08-936-165A-224/c  
; Sequence 224, Application US/08936165A  
; Patent No. 6348582

## GENERAL INFORMATION:

; APPLICANT: Black, Michael  
; APPLICANT: Burnham, Martin  
; APPLICANT: Hodgson, John

DB 2649 A 2649

## RESULT 12

US-08-842-341-1  
; Sequence 1, Application US/08842341  
; Patent No. 5953313

## GENERAL INFORMATION:

; APPLICANT: van der Bruggen, Pierre; Mandruzzato, Susanna;  
; APPLICANT: Boon-Falleur, Thierry  
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding A Functional  
; TITLE OF INVENTION: Protein, The Protein So Encoded And Uses Thereof  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; ZIP: 10022

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/842,341  
; FILING DATE: 24-APR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/718,964  
; FILING DATE: 26-September-1996  
; APPLICATION NUMBER: 08/669,590  
; FILING DATE: 24-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5953313man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5442.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2649 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

## US-08-842-341-1

Query Match 6.3%; Score 39.4; DB 2; Length 2649;  
Best Local Similarity 57.9%; Pred. No. 0.31;  
Matches 70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 495 ATTATTGAAGATACCTCGAAGAAATAGAACAAATGTAAAGTATCTGCATTGATATA 554

DB 2529 ATTATTTAAATCATAGGAATTAAGTTATCTTTAAATTTAAAGTATCTTTTTCAAAAA 2588

QY 555 TATGGCCTTAGCCATTCCAAATATCTAAATGTCAACTCAAGTAAACAAAAA 614

DB 2589 CATTTTAAATAGATAAATAATATTTGATCTTAAAAA 2648

QY 615 A 615

DB 2649 A 2649



```

; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; POLYPEPTIDES AND THEIR USES
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 224:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1886 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-936-165A-224

Query Match 6.1%; Score 37.8; DB 4; Length 1886;
Best Local Similarity 44.8%; Pred. No. 0.72;
Matches 138; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 314 AAGCCAGCTACATATCTTACCAAGAAATGAAATCATATATATCATGAGGTAATTC 373
Db 1234 ACAACACCTAATGCACTGACGAGAAAGCAGGCTGCTTAAATCAATCACTCAACT 1175

Qy 374 GAGTTTAAATACATAATCTTCCAAATCAACCAATCTTTCAGCGAAGCTGGCAGATGT 433
Db 1174 TAAAGATCAAGCAATTAATCAAAATTAATCAAAACCAAAATGATCAGGTAGACACAAC 1115

Qy 434 GGAAGGTATGGGAAAGGATAAATGTTGATGGTGAGCAAGTAATCACTAAATAATGA 493
Db 1114 TACAAATCAACGGTAAATGCTATAGATAATGTTGNAGCTGAAGTAGTAATTAACCCAAA 1055

Qy 494 CATTATTGAAGATACCTTGGAAAGAAATTTAGAAACAAAATGTAAAGTATCTGCATTGTAA 553
Db 1054 GGCAATTGCAGATNTTGAAGAAAGCTGTTAAAGAAAGCAACAGCAAAATGATATAGTCT 995

Qy 554 ATATGGCCTTAGCCATTTTCCAAATATCTAAATTTGTCAACTCAAGTAAAAAAGAAAAA 613
Db 994 TGATTCACAGATNATGAGAAAGAGTTGCTTTCACAGCATTAGCTAAAGAAAAAGAAA 935

Qy 614 AAAACTCG 621
Db 934 AGCACTTG 927

```

```

RESULT 14
US-09-040-843--3
; Sequence 3, Application US/09040843
; Patent No. 6124119
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Deborah J.
; APPLICANT: Wang, Min
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Burnham, Martin
; APPLICANT: Fosberry, Andrew
; APPLICANT: Hodgson, John E.
; APPLICANT: Lawlor, Elizabeth
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: MecB
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,843
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/057,535
; FILING DATE: 29-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10082
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2019 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-040-843-3

Query Match 6.0%; Score 37.6; DB 3; Length 2019;
Best Local Similarity 53.4%; Pred. No. 0.83;
Matches 79; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 475 AAGTAATCACTAAATAATGACATTTATGAAGATACCTTCGAAAGAAATTAGAAACAAATGT 534
Db 764 AAGTAAGCTTTAAGAGTCATACGACACCTAATAATTTAAAGAAATTTGAACAAGAAATTTG 823

Qy 535 AAGTATCTGCTATTTGATAAATATGCGCTTAGCCATTTAGCCAAATATCTAAATTTGTCACTC 594
Db 824 AAAAGTTAAAAATGAAAAAGATGCGCAGTACATGCTCAAGAGTTTCAAAATGCTGCTA 883

Qy 595 AAGTAAAAAAGAAAAAAGAAAAAAGTCA 622
Db 884 ACCTGCGTGATATAAACAACAAACTTGA 911

RESULT 15
US-09-621-855-3
; Sequence 3, Application US/09621855

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:27:34 ; Search time 1099.55 Seconds  
(without alignments)  
7647.311 Million cell updates/sec

Title: US-09-881-556-4

Perfect score: 623

Sequence: 1 cctcaaatggtgaactca.....aaaaaaaaaaactcgag 623

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	361.8	58.1	837	12	AQ411949
2	60.8	9.8	1101	12	CNS003B0
3	60.4	9.7	813	12	CNS03CIW
c 4	59.4	9.5	897	12	CNS07ABZ
5	59.2	9.5	1007	12	CNS06X9S
c 6	58	9.3	725	12	BH180166
c 7	57.2	9.2	891	12	CNS009JU
c 8	57.2	9.2	1101	12	CNS00LTJ
9	56.8	9.1	1001	12	CNS0155H
c 10	56.2	9.0	784	9	AL525973
11	55.8	9.0	1013	12	CNS02PVS
12	55.6	8.9	576	12	CNS007G4
c 13	55.6	8.9	1101	12	CNS017V2
14	55.4	8.9	430	12	CNS008K5
15	55.2	8.9	581	12	CNS01UP2
16	55.2	8.9	781	12	CNS009DO
17	55.2	8.9	996	12	CNS000F0H

c 18	55	8.8	931	12	CNS03XCF	AL264840 Tetraodon
c 19	55	8.8	982	12	AQ325799	AQ325799 nbx0021B
c 20	55	8.8	1101	12	CNS0100X	AL098379 Drosophill
c 21	55	8.8	1135	12	CNS033GQ	AL226115 Tetraodon
c 22	54.6	8.8	959	12	CNS00655	AL062806 Drosophill
c 23	54.2	8.7	395	12	CNS021CB	AL198740 Tetraodon
c 24	54.2	8.7	879	12	CNS01JRG	AL147405 Anopheles
c 25	54	8.7	828	12	CNS01ITX	AL100719 Drosophill
c 26	53.8	8.6	802	12	AZ197934	AZ197934 SP_1036_B
c 27	53.8	8.6	825	3	B1644509	B1644509 OP2248 M1
c 28	53.8	8.6	942	10	BM416122	BM416122 OP21208 M
c 29	53.8	8.6	991	12	CNS00JW8	AL077118 Drosophill
c 30	53.8	8.6	1101	12	CNS00Z15	AL097091 Drosophill
c 31	53.6	8.6	796	12	CNS0118D	AL099943 Drosophill
c 32	53.4	8.6	614	12	CNS0152H	AL104915 Drosophill
c 33	53.4	8.6	737	9	AI525942	AI525942 DU145-1.G
c 34	53.4	8.6	807	12	CNS040I3	AL300036 Tetraodon
c 35	53.4	8.6	1043	12	CNS0145P	AL103735 Drosophill
c 36	53.4	8.6	1203	12	CNS015WU	AL106008 Drosophill
c 37	53.4	8.6	1380	10	BG563591	BG563591 602582321
c 38	53.2	8.5	881	10	BF265362	BF265362 HV_C8A001
c 39	53	8.5	900	12	CNS01574	AL105514 Drosophill
c 40	53	8.5	1101	12	CNS01844	AL108862 Drosophill
41	52.8	8.5	855	12	AZ183849	AZ183849 SP_1002_A
42	52.6	8.4	811	12	CNS007W9	AL050915 Drosophill
c 43	52.6	8.4	867	12	CNS0054A	AL037618 Drosophill
c 44	52.6	8.4	870	12	AQ866797	AQ866797 nbhe0029E
c 45	52.6	8.4	871	12	AZ671726	AZ671726 ENTHP05TF

#### ALIGNMENTS

RESULT 1  
AQ411949 837 bp DNA linear GSS 12-MAY-2000  
LOCUS CPG0925A CplOWAGDNA1 Cryptosporidium parvum genomic, DNA sequence.  
DEFINITION  
ACCESSION AQ411949  
VERSION AQ411949.1 GI:4469673  
KEYWORDS GSS.  
SOURCE Cryptosporidium parvum.  
ORGANISM Cryptosporidium parvum.  
REFERENCE  
AUTHORS Strong, W.B. and Nelson, R.G.  
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis  
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)  
MEDLINE 20183851  
COMMENT Contact: Nelson, R. G.  
Depts. of Medicine & Pharmaceutical Chemistry  
San Francisco General Hospital-University of California, San Francisco  
Box 0811, San Francisco, CA 94143-0811, USA  
Tel: 415 206 8846  
Fax: 415 206 3353  
Email: malariad@itsa.ucsf.edu  
For Annotation Data see http://medasfgh.ucsf.edu/ld/CpTags/home.html  
Seq primer: M13(-21) forward  
Class: shotgun.  
Location/Qualifiers  
1. 837  
/organism="Cryptosporidium parvum"  
/strain="IOWA"  
/db\_xref="taxon:5807"  
/clone\_lib="CplOWAGDNA1"  
/lab\_host="E. coli XL2 Blue MRF"  
/note="Vector: pBlueScript II (SK-); Site 1: EcoRV; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Ivonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(<http://sequence-www.stanford.edu/group/techdev/shear.htm>). The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was subcloned into an EcoR V-digested, alkaline phosphatase-treated pBluescript II (SK-) vector and transformed into E. coli strain XL2 Blue MRF'. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

BASE COUNT 333 a 128 c 124 g 252 t  
ORIGIN  
Query Match 58.1%; Score 361.8; DB 12; Length 837;  
Best Local Similarity 98.8%; Pred. No. 4.5e-49;  
Matches 396; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 205 CTCTGCTACTATACAGAAACAGGATCCAGCTCAGATCACACT-CTGCTACTTCTCCAG 263  
|||||  
Db 13 CTCTGCTACTATACAGAAACAGGATCCAGCTCAGATCACACTCTGCTACTTCTCCAG 72  
QY 264 AAGAAGGATTGGACTCAGAACG-TTACCAATCACTTCTACGAAC-AAACTCAAAAGCCAG 321  
|||||  
Db 73 AAGAAGGATTGGACTCAGAACGCTTACCAATCACTTCTACGAACAAAACTCAAAAGCCAG 132  
QY 322 CTACATATCTTACCCAGAAAATGAATCATATAATCAGGAAGTAAATTCGAGTTTGA 381  
|||||  
Db 133 CTACATATCTTACCCAGAAAATGAATCATATAATCAGGAAGTAAATTCGAGTTTGA 192  
QY 382 ATACACTTAATCTTCAAAATCAACCAATCTTTCACGCAAGCTGCAGATGTGGAAGTT 441  
|||||  
Db 193 ATACACTTAATCTTCAAAATCAACCAATCTTTCACGCAAGCTGCAGATGTGGAAGTT 252  
QY 442 ATGGGAAAAGGATAAAATGGTTGATGGTGCAGCAAGTAACTAAAAATGACATTATG 501  
|||||  
Db 253 ATGGGAAAAGGATAAAATGGTTGATGGTGCAGCAAGTAACTAAAAATGACATTATG 312  
QY 502 AGATACCTTCGAAGAAATTAAGAACAAATGTAAGTATCTGCATTGTAATAATATGCCC 561  
|||||  
Db 313 AGATACCTTCGAAGAAATTAAGAACAAATGTAAGTATCTGCATTGTAATAATATGCCC 372  
QY 562 TTAGCCATTTCCTCAATATCTAAATTTGTCACACTCAAGTAAAA 602  
|||||  
Db 373 TTAGCCATTTCCTCAATATCTAAATTTGTCACACTCAAGTAAAA 413  
|||||

RESULT 2  
CNS003B0  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BAC08E08 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AL064078.1 GI:4941834  
fruit fly.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr)  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osoegawa and  
Aaron Mammosier in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's  
p1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.bufile.edu/drosophila\\_bac.htm](http://bacpac.med.bufile.edu/drosophila_bac.htm).

#### FEATURES

source  
1..1101  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone\_lib="RPCI-98"  
/clone="BAC08E08"  
/note="end : TET3"  
BASE COUNT 553 a 143 c 100 g 134 t 171 others  
ORIGIN

Query Match 9.8%; Score 60.8; DB 12; Length 1101;  
Best Local Similarity 38.3%; Pred. NO. 0.75;  
Matches 159; Conservative 42; Mismatches 214; Indels 0; Gaps 0;  
QY 203 ATCTTCTCTACTATACAGAAACAGGATCCAGCTCAGATCACACTCTGCTACTTCTCCA 262  
|||||  
Db 451 ANKKKKKBYTTTTTCCCNAAAYBYBYTCCCCCAAAAGAAAGCAAAACCCMA 510  
QY 263 GAAGAAGGATTGGACTCAGAACGTTTACCAATCACTTCTACAGACAACTCAAGCCAGC 322  
|||||  
Db 511 AAAAAAAMAMAMAMMCMCCCAACCAAAATACAAAMAAACCCCAAAACMAA 570  
QY 323 TACATATCTTACCAAGAAATGAATAATCAATATCAGGAAGTAAATTCGAGTTTAA 382  
|||||  
Db 571 AAAAAAMMCCCCCMMMAAAMMAAAMMAAAMMAAATTCANAAAAAACMCCCCCMA 630  
QY 383 TACACTAATCTTCCAAATCAACCAATCTTTCACGCAAGCTGGCAGATGTGGAAGTTA 442  
|||||  
Db 631 CAAAAAATTTAAAAAAMAMCMCCCMCCCMCCCMCCCMCCCMCCCMCCCMCCCMCCCMCC 690  
QY 443 TGGGAAAAGGATAAAATGGTTGATGGTGCAGCAAGTAACTAAAAATGACATTATGGA 502  
|||||  
Db 691 WAAANNAAAAAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA 750  
QY 503 AGATACCTTCGAAGAAATTAAGAACAAATGTAAGTATCTGCATTGTAATAATATGCGCT 562  
|||||  
Db 751 AAAAAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA 810  
QY 563 TAGCCATTTCCTCAATATCTAAATTTGTCACACTCAAGTAAAAAAMMAAAMMAA 617  
|||||  
Db 811 WAAATTTTAAAAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA 865  
|||||

RESULT 3  
CNS03CIW  
LOCUS  
DEFINITION  
Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone  
04B14 of library G from Tetraodon nigroviridis, genomic survey  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AL237857.1 GI:7896992  
GSS: genome survey sequence.  
Tetraodon nigroviridis.  
Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodon.  
1 (bases 1 to 813)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Roest-Crollius.H., Jaillon.O., Dasilva.C., Fizames.C., Fisher.C.,  
Bouneau.L., Billault.A., Quetier.F., Saurin.W., Bernot.A. and  
Weissenbach.J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Unpublished  
2 (bases 1 to 813)

**AUTHORS** Roest-Crollius, H., Jaillon, O., Dastilva, C., Bounneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottilier, P., Quetier, F., Saurin, W., and Weissenbach, J.

**TITLE** Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

**JOURNAL** Unpublished

**REFERENCE** 3 (bases 1 to 813)

**AUTHORS** Genoscope.

**TITLE** Direct Submission

**JOURNAL** Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

**COMMENT** This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

# FEATURES

**source** Location/Qualifiers

1. .813

/organism="Tetraodon nigroviridis"

/db\_xref="taxon:99883"

/clone\_lib="G"

/note="Genoscope sequence ID : C0B6014DA07SP1-end : PUC-ori"

**BASE COUNT** 350 a 152 c 149 g 119 t 43 others

**ORIGIN**

Query Match 9.7%; Score 60.4; DB 12; Length 813;  
Best Local Similarity 44.5%; Pred. No. 0.97;  
Matches 137; Conservative 22; Mismatches 149; Indels 0; Gaps 0;

QY 309 AACTCAAGCCAGCTACATCTCTCAACCAAGAAATCAAAATCATATATATCAGGAGGT 368  
Db 128 MNNMMVMVMMTCMTAAAGATATAAAAAAATAAAAAAATAAAAAAATAAAAAA 187  
QY 369 AATTCGAGTTTATACACTAAATCTTCCAAATCAACCCCAATCTTTCACGCAAGCTGGCA 428  
Db 188 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 247  
QY 429 GATGTGGAAGTATGGGGAAGGATAAAATGTTGATGTGGAGCAAGTAATCACTAAA 488  
Db 248 TAAGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 307  
QY 489 AATGACATATTGAGATCTCTCGAAGAAATAGAACCAAAATGTAAGCTATCTGCATT 548  
Db 308 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAA 367  
QY 549 GATAATATGGCCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAAAAAATA 608  
Db 368 CATAAATATAATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAA 427  
QY 609 AAAAAAA 616  
Db 428 AAAAAAA 435

**RESULT 4**

**CNS07ABZ/c** 897 bp DNA linear GSS 08-JUL-2001

**LOCUS** T7 end of clone BC0AA003A12 of library BC0AA from strain CBS 767 of

**DEFINITION** Debaromyomycetes hansenii, genomic survey sequence.

**ACCESSION** AL436389

**VERSION** AL436389.1 GI:12219802

**KEYWORDS** GSS.

**SOURCE** Debaromyomycetes hansenii.

**ORGANISM** Debaromyomycetes hansenii.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Debaromyomycetes.

**REFERENCE** 1 (bases 1 to 897)

**AUTHORS** Lepingle, A., Casaregola, S., Neuveglise, C., Bon, E., Nguyen, H., Artiguenave, F., Wincker, P., and Gaillardin, C.

**TITLE** Genomic exploration of the hemiascomycetous yeasts: 14.

**JOURNAL** Debaromyomycetes hansenii var. hansenii

**MEDLINE** FEBS Lett. 487 (1), 82-86 (2000)  
20584724

**REFERENCE** 2 (bases 1 to 897)

**AUTHORS** Souciet, J.-L., Aigle, M., Artiguenave, F., Blandin, G., Bolotin-Fukuhara, M., Bon, E., Brottilier, P., Casaregola, S., de-Montigny, J., Dujon, B., Durres, P., Lepingle, A., Llorente, B., Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S., Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Wincker, P., and Weissenbach, J.

**TITLE** Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

**JOURNAL** FEBS Lett. 487 (1), 3-12 (2000)

**MEDLINE** 20584711

**REFERENCE** 3 (bases 1 to 897)

**AUTHORS** Genoscope.

**TITLE** Direct Submission

**JOURNAL** Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

**COMMENT** This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces marxianus var. marxianus, Pichia lactis var. lactis, Kluyveromyces marxianus var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

# FEATURES

**source** Location/Qualifiers

1. .897

/organism="Debaromyomycetes hansenii"

/strain="CBS 767"

/variety="hansenii"

/db\_xref="taxon:4959"

/clone="BC0AA003A12"

/clone\_lib="BC0AA"

/note="end : T7"

**BASE COUNT** 148 a 70 c 47 g 590 t 42 others

**ORIGIN**

Query Match 9.5%; Score 59.4; DB 12; Length 897;  
Best Local Similarity 45.9%; Pred. No. 1.4;  
Matches 150; Conservative 13; Mismatches 164; Indels 0; Gaps 0;

QY 291 AATCATTCTACAGAACAACTCAAGCCAGCTACATATCTCAACCAAGAAATGAAAT 350  
Db 697 AATCTATTTTATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 638  
QY 351 CATAAATATCAGGAGGTAAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCAAT 410  
Db 637 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 578  
QY 411 CTTTCAGCAAGCTGGCAGATGTGGAAGTTATGGGAAAGGATAAATGGTTGGTGGT 470  
Db 577 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 518  
QY 471 GAGCAAGTAACTAAATATGACATTAATTTGAAGATACATCTCGAAGAAATTTAGAACAAA 530  
Db 517 WAAAAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 458  
QY 531 ATGTAAGTATCTGCATTGTGATAATATGCGCTTTAGCCATTTCCCAATATCTAAATTTCTCA 590  
Db 457 AATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 398  
QY 591 ACTCAAGTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 617  
Db 397 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 371

**RESULT 5**

**CNS06X9S**

**LOCUS** T3 end of clone AX0AA039F08 of library AX0AA from strain CBS 7064

**DEFINITION** of Pichia farinosa, genomic survey sequence.

**CNS06X9S** 1007 bp DNA linear GSS 06-JUL-2001

[illegible]

QY	332	TAACCAAGAAATGAAATCATATATATCAGGAGGTAAATTCGAGTTTTTAATACATAAA	391
Db	672	TWATAATAAARAATBAAAAAAMAAAAAAMAAAAAAMAAAAAAGSTSBYSBTNSAVA	613
QY	392	TCITCCAAATCAACCCCAATCTTTCACGCAAGCTGGCAGATGTGCAAGATTATGGGGA	451
Db	612	TSWAABBAAMAAAAAAYSBYTATTSBBBAAMAAAAAAGAKGKAAGRAAGAAYA	553
QY	452	GGATAAATGGTGTGAGCAAGTAATCACTATAAATGCATATATGCAAGATACCTC	511
Db	552	AGAAGAGGGRAMAMCTYACMCCTCRYTAMAAATTAACACTTAATATATAAATAATTA	493
QY	512	GAAGAATAATTAGAACAAAAATGTAAAGTATCTGCATGTATAATATCGCCTTACCCAT	571
Db	492	AWATTAAAAAAMAAAAAATCTTTHAAATCYWAAAAAAMAAAAATCAATTTTAAAYAY	433
QY	572	CCAAATATCAATTTGTCAACTCAAGTAAAAAAMAAAAAAMAAAAA 617	
Db	432	TYTAHAATAATYAAAAAAMAAAAAHTTAATAAAMAAAAAAMAAAAA 387	
RESULT 8			
CNS00LT2/c			
LOCUS			
DEFINITION			
Drosophila melanogaster genome survey sequence FET3 end of BAC:			
BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit			
fly), genomic survey sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Determination of this BAC-end sequence was carried out as part of a			
collaboration with the Berkeley Drosophila Genome Project (BDGP).			
The BDGP is constructing a physical map of the Drosophila			
melanogaster genome using these BACs. For further information			
please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila			
melanogaster BAC library was prepared by Kazutoyo Osoegawa and			
Aaron Mammos in Pieter de Jong's laboratory in the Department of			
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,			
NY. The library is named RPCI-98 and was constructed by partial			
EcoRI digestion of Drosophila DNA provided by the BDGP from the			
isogenic strain y2; cn bw sp, the same strain used for the BDGP's			
P1 and EST libraries. A more detailed description of the library			
and how to order individual BAC clones, the entire library, or			
filters for hybridization from the BACPAC Resource Center can be			
found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .			
FEATURES			
source			
1. .1101			
/organism="Drosophila melanogaster"			
/db_xref="taxon:7227"			
/clone_lib="RPCI-98"			
/clone="BACR48P19"			
/note="end : TET3"			
BASE COUNT			
ORIGIN			
469 a 6 c 151 t 406 others			
Query Match 9.2%; Score 57.2; DB 12; Length 1101;			
Best Local Similarity 19.9%; Pred. No. 2.8;			
Matches 116; Conservative 193; Mismatches 273; Indels 0; Gaps 0;			
QY	36	CAAGCCCTATTACTCACCGAAGAAAGCAATCAAGTCAATCATCTTCTGTGACAGTT	95
Db	1072	MMMMMAAMMMCMCMCCCKCHMMCAAMMMHMTTMMKMMKMMMAAMMMMAAMMM 1013	







Matches 147; Conservative 34; Mismatches 163; Indels 4; Gaps 1;

QY	271	ATTGACTCAGAAAGCTTTACCAATCACCTTCTACAGAACAAACTCAAAGCCAGCTACATATC	330
Db	189	AWTAGGCTAAGAAAAAATAA	248
QY	331	CTRACCAAGAAATGAAAATCATATAATATCAGGAAAGTAAATTCAGTGTTTAATACACTAA	390
Db	249	RCACGCGCGCAAAWAAAAACAAAAAATAATATAAAAAAAAAAAAAAAAAAAAAAAWAA	308
QY	391	ATCTTCCAAATCAACCCCAATCTTTTCACGCAAGCTGGCAGATGTGGAAGTTATGGGAAA	450
Db	309	GWAAAAAATAAAACCGACCCACCAWATATWAAAAATACGAATCGGTAAAWATAAA	368
QY	451	AGGATAAAATGTTTGATGGTGACCAAGTAATCACTAAAAATGACATTTATTGGAAGACTTT	510
Db	369	GAAGTATTAATAG----TGATRWCGTGGAGAGAAAAAAGGGCAATAWAAAAAATWA	424
QY	511	CGAAAGAAATATAGAAACAAAAATGTAAAGTATCTGCATTTGATAAATATGGCTTAGCCATT	570
Db	425	ACAAAAAATAAAWAAACAGCTWATAAAATTTATAAAAAWAAWAAWAAACAAAAAAWA	484
QY	571	TCCAAATATCTAAATTTGCTACCTCAAGTAAAAAATAAAAAAAAAAAAAAAAC	618
Db	485	CAAAAAAATAAAWAAAAAATAAAWAAAAAAGAAAAAATAAAAAAAAAAAAAAAC	532

RESULT 13

CONS017v2/c

LOCUS

DEFINITION

CONS017V2 1101 bp DNA linear GSS 26-JUL-1999

Drosophila melanogaster genome survey sequence sp6 end of BAC

BAC37L10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION	AL108536	
VERSION	AL108536.1	GI:5628840
KEYWORDS	GSS:	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS	1 (bases 1 to 1101)	
TITLE	Genoscope.	
JOURNAL	Direct Submission	
COMMENT	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)	
	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.	

FEATURES	source	Location/Qualifiers
BASE COUNT	268 a	174 c 114 g
ORGANISM		1. 1101
		/organism="Drosophila melanogaster"
		/plasmid="pBelOBAC11"
		/db_xref="taxon:7927"
		/clone_lib="DrosBAC"
		/clone="BACN37L10"
		/note="end : Srf"
	268 a	174 c 114 g 366 t 179 others

Query Match	8.9%;	Score 55.6;	DB 12;	Length 1101;
Best Local Similarity	40.9%;	Pred. No. 5.1;		
Matches 130;	Conservative 33;	Mismatches 155;	Indels 0;	Gaps 0;
QY	300	TACAGAACAACTCAAGCCAGCTACATATCTTACCGAGAAATGAAATCATATAT	359	
	:   :	:   :   :		
Db	1093	TATWAAWAAATATWAHATWAAWAAATATWTWTWTAAGAAAAWAAAAA	1034	

Qy	360	CAGGAAGGTAAATCGAGCTTTTAATACACTAAATCTTCCAATCAACCCAAATCTTTCACGC	419
Db	1033	AAAAAAAAAHHAWAAAAATTTTAAMWTAATAAAAATHATAAAWWAAAAAWATAATTTTTT	974
Qy	420	AAGCTGCCAGATGCGGAAAGTTATGGGGAAGAAGCATAAATGGTGTGATGCGACRAGTA	479
Db	973	AAAYAAAYAAATAAWAAANNAANAANTWAANAANAANAANAANAATAATAAAAAAAA	914
Qy	480	ATCAGTAAATGACATTATTGGAAGTACTTCGAAAGAAATTAGAAAAACAATGTAAGT	539
Db	913	AA	854
Qy	540	ATCTGCATTCATAAATATGCCITAGCCATTCCAAATATCTAAATGTGCTCACTCAGTA	599
Db	853	AAAAATAAWAAAAATATAAYTAANAANAANAANAANAANAANAANAANAANAANA	794
Qy	600	AAAAAAAAAAAAAAAAAAAA	617
Db	793	TAAWAAAAAAAAAAAAAAAA	776

RESULT 14  
 CNS0008K5  
 LOCUS  
 DEFINITION  
 430 bp DNA linear GSS 03-JUN-1999  
 Drosophila melanogaster genome survey sequence TET3 end of BAC #  
 BACR17M14 of RPC1-98 library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.  
 ACCESSION  
 AL052179  
 VERSION  
 AL052179.1 GI:4933829  
 KEYWORDS  
 GSS.  
 SOURCE  
 fruit fly.  
 ORGANISM  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 430)  
 Genoscope.  
 REFERENCE  
 AUTHORS  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial *EcoRI* digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and Est libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```

FEATURES
  source
    Location/Qualifiers
      1. .430
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
        /clone_lib="RPCI-98"
        /clone="BACRI7M14"
        /note="end : TET3"
      331 a
        27 c 18 g 10 t 44 others
BASE COUNT
ORIGIN

```

```
Query Match      8.9%; Score 55.4; DB 12; Length 430;
Best Local Similarity 42.7%; Pred. No. 7.6;
Matches 146; Conservative 17; Mismatches 179; Indels 0; Gaps 0;
QY 276 ACTCAGAAGGTTACCAATCACCTTCTACAGAACAAACTCAAAGCCAGCTACATATCTCAAC 335
```

```
Db 7 ACAGAGAAATGAANNACNNACNNNNNAACCCCAANNANACANACACMCAAAATCAAGA 66
QY 336 CAAGAAATGAAATCATATAATATCAGGAAGGTAATTCGAGTTTATATACACATAAATCTT 395
Db 67 AAAAAAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 126
QY 396 CCAATCAACCCCAATCTTTTCAGCAAGCTGGCAGATGTGGAAGTTATGGGAAAGGAT 455
Db 127 AARAAACCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 186
QY 456 AAAATGTTGATGTGAGCAAGTAATCACTAAATAATGACATATTATGAAGTACTTCCGAA 515
Db 187 ATAATAATTAAGGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 246
QY 516 GAAATTAGAAACAAATGTAAGTATCTGCATGATAAATATATGCGCTTACCCATTTCCAA 575
Db 247 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 306
QY 576 ATATCTAAATTCCTCACTCAAGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 617
Db 307 MAATAAATAAAGMCMCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 348

RESULT 15
CNS01UPZ
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
197M17 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION
AL168128
VERSION
AL168128.1 GI:7806185
KEYWORDS
GSS; genome survey sequence.
SOURCE
Tetraodon nigroviridis.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE
1 (bases 1 to 581)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 581)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 581)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
Location/Qualifiers
1. .581
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="197M17"
/clone_lib="G"
/note="Genoscope sequence ID : COAG197AG09SP1-end :
PUC-ori"
BASE COUNT 417 a 7 c 58 g 31 t 68 others
ORIGIN
Query Match 8.98; Score 55.2; DB 12; Length 581;
Best Local Similarity 39.68; Pred. No. 7.4;
```

```
Matches 141; Conservative 23; Mismatches 192; Indels 0; Gaps 0;
QY 262 AGAAGAGGATTGGACTCAGAACGTTTACCATCCTTCTACAGAACAACTCAAGCCAG 321
Db 212 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 271
QY 322 CTACATATCCTAACCAAGAAATGAAATCATATAATCATCAGGAAGGTAAATTCGAGTTT 381
Db 272 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 331
QY 382 ATACACTAAATCTTCCAAATCAACCCCAATCTTTCACGCAAGCTGGCAGATGTGGAAG 441
Db 332 AATCAATAATAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAWAW 391
QY 442 ATGGGGAAAGGATAAAATGTTGATGGTGAGCAAGTAATCACTTAAATAATGACATTT 501
Db 392 AAAAAAAAAAWATTAAGAAAGAAWAGAGGAAAGGAGGAAAAATATAATATAAATAA 451
QY 502 AAGATACTTCCAAAGAAATTTAGAAACAAATGTAAAGTATCTCCATTTGATTAATATG 561
Db 452 AAGGAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 511
QY 562 TTAGCCATTTCCAAATATCTTAAATTTGCAACTCAAGTAAAAAATAAATAAATAAATA 617
Db 512 TAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 567

Search completed: November 2, 2002, 01:47:26
Job time : 1112.55 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:16:04 ; Search time 2065.61 Seconds  
(without alignments)  
7331.704 Million cell updates/sec

Title: us-09-881-556-3

Perfect score: 700

Sequence: 1 gaattggcagcagattttt.....accgcctctacagaaagc 700

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/1/pna/PCTUS\_COMB.seq:\*

2: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq:\*

3: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq:\*

4: /cgn2\_6/ptodata/1/pna/US080\_COMB.seq:\*

5: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq:\*

6: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq:\*

7: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq:\*

8: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq:\*

9: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq:\*

10: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq:\*

11: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq:\*

12: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq:\*

13: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq:\*

14: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq:\*

15: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq:\*

16: /cgn2\_6/ptodata/1/pna/US092\_COMB.seq:\*

17: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq:\*

18: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:\*

19: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq:\*

20: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq:\*

21: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq:\*

22: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq:\*

23: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq:\*

24: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq:\*

25: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq:\*

26: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq:\*

27: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq:\*

28: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq:\*

29: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq:\*

30: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq:\*

31: /cgn2\_6/ptodata/1/pna/US098A\_COMB.seq:\*

32: /cgn2\_6/ptodata/1/pna/US098B\_COMB.seq:\*

33: /cgn2\_6/ptodata/1/pna/US098C\_COMB.seq:\*

34: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq:\*

35: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq:\*

36: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq:\*

37: /cgn2\_6/ptodata/1/pna/US100\_COMB.seq:\*

38: /cgn2\_6/ptodata/1/pna/US101\_COMB.seq:\*

39: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq:\*

40: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq:\*

41: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq:\*

42: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq:\*

43: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq:\*

44: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq:\*

45: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq:\*

46: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq:\*

47: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq:\*

48: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq:\*

49: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq:\*

50: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq:\*

51: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq:\*

52: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq:\*

53: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq:\*

54: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq:\*

55: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq:\*

56: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq:\*

57: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq:\*

58: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq:\*

59: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq:\*

60: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq:\*

61: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq:\*

62: /cgn2\_6/ptodata/1/pna/US6023\_COMB.seq:\*

63: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq:\*

64: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq:\*

65: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq:\*

66: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq:\*

67: /cgn2\_6/ptodata/1/pna/US6028\_COMB.seq:\*

68: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq:\*

69: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq:\*

70: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq:\*

71: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq:\*

72: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq:\*

73: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq:\*

74: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq:\*

75: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	700	100.0	700	33	US-09-881-556-3
2	54.8	7.8	342	33	US-09-865-419A-12796
3	53.6	7.7	481	31	US-09-823-301-7836
4	52.6	7.5	236	26	US-09-669-817A-17885
5	51	7.3	362	33	US-09-865-419A-5928
6	50.2	7.2	252	17	US-09-388-907-6637
7	50	7.1	445	26	US-09-666-355A-5285
8	50	7.1	537	37	US-10-021-323-7228
9	50	7.1	537	64	US-60-255-619-7228
10	49.2	7.0	46657	64	US-60-258-273-22
11	48.6	6.9	419	18	US-09-480-903-11234
12	48.6	6.9	419	36	US-09-960-352-11234
13	48.4	6.9	423	17	US-09-394-745-4413
14	48.4	6.9	423	22	US-09-565-306-68010
15	48.4	6.9	423	25	US-09-654-617-330054
16	48.4	6.9	423	27	US-09-684-016-330054
17	48.2	6.9	552	58	US-60-197-873-63174
18	48.2	6.9	1141	31	US-09-806-708A-22
19	48	6.9	412	58	US-60-197-872-74895
20	48	6.9	2000	33	US-09-887-272A-5263
21	47.8	6.8	445	19	US-09-521-640-237677
22	47.8	6.8	445	33	US-60-140-769-32426
23	47.8	6.8	541	33	US-09-865-439A-110751
24	47.6	6.8	605	7	US-08-358-918-36
25	47.6	6.8	832	24	US-09-621-976-2813
26	47.6	6.8	832	53	US-60-147-499-2813
27	46.6	6.7	399	29	US-09-758-109-12204
28	46.6	6.7	874	29	US-09-726-802-1504
29	46.6	6.7	978	19	US-09-521-640-274601
30	46.4	6.6	510	59	US-60-207-458-49190
31	46.4	6.6	399	64	US-60-253-652-10503
					Sequence 3, Appli
					Sequence 12796, A
					Sequence 7836, Ap
					Sequence 17885, A
					Sequence 5928, Ap
					Sequence 6627, Ap
					Sequence 5285, Ap
					Sequence 7228, Ap
					Sequence 22, Appl
					Sequence 11234, A
					Sequence 11234, A
					Sequence 44113, A
					Sequence 68010, A
					Sequence 330054,
					Sequence 330054,
					Sequence 63174, A
					Sequence 22, Appl
					Sequence 74895, A
					Sequence 5263, Ap
					Sequence 237677,
					Sequence 32426, A
					Sequence 110751,
					Sequence 36, Appl
					Sequence 2813, Ap
					Sequence 12204, A
					Sequence 1504, Ap
					Sequence 274601,
					Sequence 49190, A
					Sequence 10503, A

C 32 46.2 6.6 418 17 US-09-304-517A-259577 Sequence 259577,  
C 33 46.2 6.6 418 36 US-09-985-678-259577 Sequence 259577,  
C 34 46.2 6.6 446 17 US-09-371-146A-259449 Sequence 259449,  
C 35 46.2 6.6 446 22 US-09-565-240-8353 Sequence 8353, Ap  
C 36 46.2 6.6 446 25 US-09-654-617-67534 Sequence 67534, A  
C 37 46.2 6.6 446 27 US-09-684-016-67534 Sequence 67534, A  
C 38 46.2 6.6 446 27 US-09-684-016-67534 Sequence 67534, A  
C 39 45.8 6.5 408 33 US-09-865-419A-1419 Sequence 53836, A  
C 40 45.6 6.5 319 25 US-09-654-617-317618 Sequence 317618, A  
C 41 45.6 6.5 319 27 US-09-684-016-317618 Sequence 317618, A  
C 42 45.6 6.5 339 17 US-09-394-745-12853 Sequence 12853, A  
C 43 45.6 6.5 339 22 US-09-565-306-33296 Sequence 33296, A  
C 44 45.6 6.5 360 58 US-09-197-872-11831 Sequence 11831, A  
C 45 45.6 6.5 565 31 US-09-824-130-4331 Sequence 4331, Ap

## ALIGNMENTS

RESULT 1  
US-09-881-556-3  
; Sequence 3, Application US/09881556  
; GENERAL INFORMATION:  
; APPLICANT: Simonsen, J. Neil  
; TITLE OF INVENTION: Cryptosporidium Parvum Antigens, Antibodies Thereof and  
; Diagnostic and Therapeutic Compositions Thereof  
; FILE REFERENCE: 9000-0054  
; CURRENT APPLICATION NUMBER: US/09/881,556  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/212,083  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 700  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:C. Parvum AG2  
US-09-881-556-3

Query Match 100.0%; Score 700; DB 33; Length 700;  
Best Local Similarity 100.0%; Pred. No. 9.6e-116;  
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATTCGGCAGAGATTTTCTTTTACCTATTTCAATTTAGTTCTTTGATTCAA 60  
DB 1 GAATTCGGCAGAGATTTTCTTTTACCTATTTCAATTTAGTTCTTTGATTCAA 60  
QY 61 ACGATGCAAGTCATTTTGTAAATCCAGATGGATCCGGAATTTTGAAACATTT 120  
DB 61 ACGATGCAAGTCATTTTGTAAATCCAGATGGATCCGGAATTTTGAAACATTT 120  
QY 121 CTACTAAATTCGAATTTAAATTTGAGTTCGCTTCAATTTGAGTTCGCGAGTG 180  
DB 121 CTACTAAATTCGAATTTAAATTTGAGTTCGCTTCAATTTGAGTTCGCGAGTG 180  
QY 181 ATATTTTATCTTGATAGGAACACGCTCTTGAAGCTGTAGTTTATCAATCGCTTG 240  
DB 181 ATATTTTATCTTGATAGGAACACGCTCTTGAAGCTGTAGTTTATCAATCGCTTG 240  
QY 241 TTTTCTATACAAAACATGTTTGAAGAAGATGAAGACATGCTTTAAACCCCTTTAATC 300  
DB 241 TTTTCTATACAAAACATGTTTGAAGAAGATGAAGACATGCTTTAAACCCCTTTAATC 300  
QY 301 GCGCTGAGAATAAATGACTTTTGGTCTTGAGAAAGACTTAGCGACAACCTCTCCAACTTT 360  
DB 301 GCGCTGAGAATAAATGACTTTTGGTCTTGAGAAAGACTTAGCGACAACCTCTCCAACTTT 360  
QY 361 CTAATTCGAATTTATCTTTTCCCTTACATGGAATAAATGATTCCTTGATATATTC 420  
DB 361 CTAATTCGAATTTATCTTTTCCCTTACATGGAATAAATGATTCCTTGATATATTC 420

QY 421 TTAACACAAATTAATCAACAAAAATTTCTCTTGAAGTTCCGAGAAAACTCAATCTCA 480  
DB 421 TTAACACAAATTAATCAACAAAAATTTCTCTTGAAGTTCCGAGAAAACTCAATCTCA 480  
QY 481 CAATTGGATATTGGAGTATTATCGATGGATTTCTCTCTTTTAATTAACATGCGCTA 540  
DB 481 CAATTGGATATTGGAGTATTATCGATGGATTTCTCTCTTTTAATTAACATGCGCTA 540  
QY 541 TAAAGAAAAATGCCACCTTGAATTAATCAAGAAATCAAAATATTCAAAATGAATTAATGAAG 600  
DB 541 TAAAGAAAAATGCCACCTTGAATTAATCAAGAAATCAAAATATTCAAAATGAATTAATGAAG 600  
QY 601 CCATTAACCTCAACAAATCCAGAAATCAGTGGGAATAACTTAACTCAGAACCAACACACA 660  
DB 601 CCATTAACCTCAACAAATCCAGAAATCAGTGGGAATAACTTAACTCAGAACCAACACACA 660  
QY 661 AAGCCTCATCCAGTTGTAGACCGCATCTACAGAAAAGC 700  
DB 661 AAGCCTCATCCAGTTGTAGACCGCATCTACAGAAAAGC 700

RESULT 2  
US-09-865-419A-12796/C  
; Sequence 12796, Application US/09865419A  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Kunsheng  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(51935)B  
; CURRENT APPLICATION NUMBER: US/09/865,419A  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: US 60/208,063  
; PRIOR FILING DATE: 2000-05-31  
; NUMBER OF SEQ ID NOS: 54020  
; SEQ ID NO 12796  
; LENGTH: 342  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3236-147-P1-N1-H5  
US-09-865-419A-12796

Query Match 7.8%; Score 54.8; DB 33; Length 342;  
Best Local Similarity 49.3%; Pred. No. 3.4;  
Matches 143; Conservative 0; Mismatches 147; Indels 0; Gaps 0;  
QY 373 ATTATCTTTTCCATACATGGAATAACTCGATTTCTGGATATATTCCTCAATTAACACATA 432  
DB 294 ATTTATATTTTAAAAAAGTAACTTTTCTTTTAAATATAAATGAATAATTTTATT 235  
QY 433 AAATCAACAAATTTCTCTTGAAGTTCCGAGAAAACTCAATCTCCACATTTGGATATT 492  
DB 234 ATACAAAAAATATTTCTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTATAT 175  
QY 493 GGAGTATTATCGATGGATTTTCTTCTTTAAATTAACATCGCGCTATAAAGAAAAATG 552  
DB 174 AATGATTTTATTTTATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTT 115  
QY 553 GCCACTTCAATTAACAAATCAAAATTAATCAAAATTAATCAAAATTAATCAAAATTAATCA 612  
DB 114 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTT 55  
QY 613 ACAATTCAGATCAGGTGGGAATAACTTAACCTCAGAACCAACCAACCAACCAACCAACCA 662  
DB 54 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTT 5

RESULT 3  
US-09-823-301-7836/C  
; Sequence 7836, Application US/09823301  
; GENERAL INFORMATION:  
; APPLICANT: Shyjan, Andrew W.

```

: APPLICANT: Macbeth, Kyle J.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.2085-001
: CURRENT APPLICATION NUMBER: US/09/823.301
: CURRENT FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,421
: PRIOR FILING DATE: 2000-03-31
: NUMBER OF SEQ ID NOS: 10002
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7836
: LENGTH: 481
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(481)
: OTHER INFORMATION: n = A,T,C or G
: US-09-823-301-7836

```

Query Match	7.7%	Score 53.6;	DB 31;	Length 481;
Best Local Similarity	43.1%	Pred. No. 5.9;		
Matches 146;	Conservative	0;	Mismatches 193;	Indels 0;
Gaps	0;			

  

QY	320	TTTTGGTTCGAGAAAGACATTACGACAACTCTCCAACTCTCTAAATTCGAAATATTATCT	379
DB	367	TTTTTNATNTTAAAAAANTATNGTTTTTTTTTAAAAATTTTTTTTACANGGGTTTT	308
QY	380	TTTCCCTTACATGGAATAACTGCAATCTTGGATATATTCCAATTAACACAAATAAAATCAA	439
DB	307	TTTTTTGAAAAAANAANAANCNNNTTTTTTTTTTGNNTTCCCCCNAAANNAANTAA	248
QY	440	CAAAATTTCTCTTGAAGTTCCGGAGAAAACTCAATCTCCACAAATGGATATTGGAGTAT	499
DB	247	AAATNTTTTTTTTTTTTTTTTANAAAAATNNTTTTTTAAAAANTTTTNAANNGNTTAA	188
QY	500	TATCGATGGATTTTCTTCTTTTAAATTAACATCGCCCTATAAAGAAAAATGCCACTT	559
DB	187	ATTNAAAAAANAATTCNCNCNTTTTTTTTTTTCOCANNATAAAAAANAANAANAANA	128
QY	560	GAATAATCAAGAATCAAAATATTCAAAATGAATTAATGAAGCCACTAAACTCAACAATC	619
DB	127	AAAAATAAAAAANNAANAANAANAANAANAANAANAANAANAANAANAANAANA	68
QY	620	CAGATCAGTGGGAATTACTTTACTCAGAACCCAAACA	658
DB	67	AAAAAAAAAANAAGNAANAANAANAANAANAANAANAANAANAANAANAANAANA	29

```

RESULT 4
US-09-669-817A-17885
: Sequence 17885, Application US/09669817A
: GENERAL INFORMATION:
: APPLICANT: Bytun, Joseph R.
: APPLICANT: Ruan, Yijun G.
: APPLICANT: Wallick, C. Kevin
: TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(51469)B
: CURRENT APPLICATION NUMBER: US/09/669,817A
: CURRENT FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: US 60/156,951
: PRIOR FILING DATE: 1999-09-30
: NUMBER OF SEQ ID NOS: 43701
: SEQ ID NO 17885
: LENGTH: 236
: TYPE: DNA
: ORGANISM: Oryza sativa
: FEATURE:
: OTHER INFORMATION: Clone ID: uC-osflcyp022e09b2
US-09-669-817A-17885

```

Query Match 7.5%; Score 52.6; DB 26; Length 236;

```

Best Local Similarity 55.1%; Pred. No. 8;
Matches 103; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 484 TTGGATATTGGAGTATTATCGATGGATTCTCTCTCTTAATTAAACGTCGCTATAA 543
Db 49 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGAAAAATTTAAAAATAAAA 108
QY 544 AACAAATGGCCACTTGGATTAATCAAGATCAAAATATTCAAAATGAAATATGAAGCCA 603
Db 109 AAAAAAGGTTAAAAAAAATAAAAAAAAATAAAAAATAATGAAAAAATAAAAAAATAAAAAA 168
QY 604 CTAAACTCAACAAATCCGAATCAGGTGGGAATAACTTTAACTCAGAACCAAAACACAAG 663
Db 169 ATAAAAAATAAAAAAAAATTTATTAAAGAAAAATAAATAAAAAAACAGGAAAAAATAAAA 228
QY 664 CCTCATC 670
Db 229 TATAATC 235

RESULT 5
US-09-865-419A-5928/c
; Sequence 5928, Application US/09865419A
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(51935)B
; CURRENT APPLICATION NUMBER: US/09/865,419A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,063
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 54020
; SEQ ID NO 5928
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(362)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3236-069-p1-N1-C7
US-09-865-419A-5928

```

```

Query Match          7.3%; Score 51; DB 33; Length 362;
Best Local Similarity 47.2%; Pred. NO. 16;
Matches 150; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 345 ACAACTCTCCAACTCTCTAATCTCGAATATTTATCTTTTCCCTTACATGGAATAACATGCAATT 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 342 AAAAATAAATAAATAATTTATTAATAAANATATTTATTAAAAATTTTAAAAATAAANAATT 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 CTTGGGATATATTCCTAATTAACACAAATAAAATCAACAAAATTTCTCTTGAAGATTCCCGA 464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 TTTTATATAATAATATTAATAAATAAATAAATAATTTTCTTTTAAAAAAAACITTTAA 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 465 GAAACTCAATCTCCACAATGAGATATCGAGATATATCGATGGAATTTCTCTCTCTTTTA 524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 AATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 525 ATTAAACATCGCCTATAAAGAAATGCCACTTCAATAATCAAGAAATCAAAATATTCA 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 TTTTANAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 585 AAATGAAATTAATGAAGCCACTAAACTCAACAATAATCCAGAAATCAGGTGGGAATTAAC 644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 43
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 645 TCAGAACCAAAACACAAA 662
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 42 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 25
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 6

US-09-388-907-6627/c

; Sequence 6627, Application US/09388907

; GENERAL INFORMATION:

; APPLICANT: Searle Monsanto

; APPLICANT: Doshi, Parul D.

; APPLICANT: Dotson, Stanton B.

; APPLICANT: Chen, Jin

; APPLICANT: Ma, Xiao-Jun

; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with HS-5

; FILE REFERENCE: SO-3145

; CURRENT APPLICATION NUMBER: US/09/388,907

; CURRENT FILING DATE: 1999-09-02

; EARLIER APPLICATION NUMBER: US 60/0009286

; EARLIER FILING DATE: 1998/09/04

; NUMBER OF SEQ ID NOS: 7558

; SEQ ID NO 6627

; LENGTH: 252

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB33-004-Q1-B1-C6

US-09-388-907-6627

## Query Match

Best Local Similarity 7.2%; Score 50.2; DB 17; Length 252;

Matches 97; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 488 ATATTGGAGTATTATCGATGGATTCTCTCTTTAAATTAACATCGCGCTATAAAGA 547

DB 188 ATTTAAAAAAATTTTGTGAATTTTAAATTTCAAAATTAATATTCCTGAAAAAAA 129

QY 548 AATGGCCACTTGAATCAAGAAATCAAAATATTCATAATGAATTAATCAAGCCACTAA 607

DB 128 AATATTCCTCTTAAAGTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 69

QY 608 ACTCAACAATCCAGATCAGGTGGGAATTAACCTTAACCTCAGAACCAACACAAA 662

DB 68 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAAAAAAAAAAAAAAAAAA 14

## RESULT 7

US-09-666-355A-5285

; Sequence 5285, Application US/09666355A

; GENERAL INFORMATION:

; APPLICANT: Fincher, Karen L.

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Xiao, Jinhua

; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With

; FILE REFERENCE: 38-21(51462)B

; CURRENT APPLICATION NUMBER: US/09/666,355A

; CURRENT FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US 60/155,005

; PRIOR FILING DATE: 1999-09-21

; NUMBER OF SEQ ID NOS: 25978

; SEQ ID NO 5285

; LENGTH: 445

; TYPE: DNA

; ORGANISM: Gossypium hirsutum

; FEATURE:

; OTHER INFORMATION: Clone ID: uC-gsronu33B169f10b1

US-09-666-355A-5285

## Query Match

Best Local Similarity 7.1%; Score 50; DB 26; Length 445;

Matches 122; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 421 TTAACACAAATAAATCAACAAAATTTCTTTGAAAGTTCGCGAGAAAATCAATCTCAA 480

DB 117 TTAATTTAGGTAAGAAAAATCAAAATGCTCAAAAGCATTTGCTAGAGATTATTGAAGTCAA 176

## Query Match

Best Local Similarity 7.1%; Score 50; DB 37; Length 537;

Matches 152; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 16 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 75

DB 71 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 130

QY 76 TATTTGTTTTTAAATCCAGATCGGATTTTGAAGAAATTTTGAAGAAATTTTCTACTAAATTCGAAA 135

DB 131 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 190

QY 136 TTAATTTTTCAGCTTCGCTTGATAAATGGTAGTGGCTCGGAGGTGATATTTTATCCPTG 195

DB 191 AAAAAATAAAACCAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 250

QY 196 ATAGAAACACGCTCTTGAAGCTGTAAGTATTCAATCGCTTGCTTTCTTCTATACAAAA 255

DB 251 TTGAGAAAGCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 310

QY 256 CATGTTTTTGAAGAAATGAAGCACAATTCCTTAAACCCCTTTAATCGCTCGAGATAAAAA 315

DB 311 TTTTCTTTTGAATAGGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 370

QY 316 TCACCTTTTGGTTCGAGAAAGA 337

DB 371 AAAAAATTTTTTTTAAGGAAAA 392

## RESULT 9

US-60-255-619-7228

; Sequence 7228, Application US/60255619







```

US-09-654-617-330054/c
; Sequence 330054, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovacic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 330054
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Zea mays
US-09-654-617-330054

Query Match          6.9%; Score 48.4; DB 25; Length 423;
Best Local Similarity 47.8%; Pred. No. 49;
Matches 172; Conservative 0; Mismatches 186; Indels 2; Gaps 1;

Qy 305 TGAGAATAAAATGACTTTTGTTCTCGAGAAAGACTTTAGCGACAACCTCTCCCAATCTTCTAA 364
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 TTAATATAAAATAAATTAATAAACCCAAAATAAAATAAAAAATAATATAATAATCCAAATAA 328

Qy 365 TTCTGAATATTATCTTTTCCTTACATCGAATACTGCATCTCTCGATATATATCCAAATAA 424
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 327 ATATAAATAAAATAAAATAAAATTTTACAAATAAAATAAATTTTAAATAAATAAAAAACAATAA 268

Qy 425 CACAATAATAAATCAACAAAATTTCTCTTTGAAGTTTCGGAGAAAACCTCAATCTCCACAAT 484
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 AAAAAATTATATAACAATAAAAAATAAATAAATAATATAATAAATAAAAAATAAAAAA 208

Qy 485 TGGATATTGGAGTATTATCGAGGATTTCTCTCTTTTAAATTAACATCGCGCTATAAA 544
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 AATATAATAATTAATAAACCTCATTTATTTATATAAAAAATAAAAAATAATTCACCTAA 148

Qy 545 AGAAATGGCCACTTGAATAA--TCAAGAACTAAAAATATTCAAAAATGAAATAATGAAGCC 602
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 ATTAATAAANAATTAATAAATAAATAATATAAATAAATAAATAAATAAATAAATAAAC 88

Qy 603 ACTAACTACAAANAATCCAGAAATCAGGTGGGAATACTTAATCTCAGAACCAACAAACAAA 662
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 87 CCAAAACCCCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 28

```

Search completed: November 2, 2002, 00:37:01  
Job time : 2081.61 secs

THIS PAGE BLANK (u8PTD)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:55:48 ; Search time 271.958 Seconds  
(without alignments)  
9300.473 Million cell updates/sec

Title: US-09-881-556-3

Perfect score: 700

Sequence: 1 gaattggcagcagattttt.....accgctctctacagaaagc 700

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2893907 seqs, 1806667767 residues

Total number of hits satisfying chosen parameters: 5787814

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_New:

- 1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq.\*
- 8: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	48.4	6.9	7040	12	US-10-172-086-13
C 2	47.2	6.7	5501	13	US-10-221-613-204
C 3	46.2	6.6	639	9	US-09-814-353-6171
C 4	46.2	6.6	639	9	US-09-814-353-12449
C 5	46.2	6.6	639	9	US-09-814-353A-6171
C 6	46.2	6.6	639	9	US-09-814-353A-12449
C 7	46	6.6	539	9	US-09-814-353-5852
C 8	46	6.6	539	9	US-09-814-353-12133
C 9	46	6.6	539	9	US-09-814-353A-5852
C 10	46	6.6	539	9	US-09-814-353A-12133
C 11	45.6	6.5	752	9	US-09-814-353-18656
C 12	45.6	6.5	752	9	US-09-814-353A-18656
C 13	45.4	6.5	874	13	US-10-198-846-5980
C 14	45.4	6.5	5680	13	US-10-221-613-3
C 15	45.2	6.5	327	9	US-09-785-276A-8158
C 16	45.2	6.5	477	9	US-09-918-995-15842
C 17	45.2	6.5	18283	13	US-10-221-613-326
C 18	45	6.4	6394	13	US-10-221-613-239

Sequence 5831, Ap	19	44.8	6.4	500	9	US-09-814-353-5831	Sequence 5831, Ap
Sequence 12112, A	20	44.8	6.4	500	9	US-09-814-353-12112	Sequence 12112, A
Sequence 5831, Ap	21	44.8	6.4	500	9	US-09-814-353A-5831	Sequence 5831, Ap
Sequence 12112, A	22	44.8	6.4	500	9	US-09-814-353A-12112	Sequence 12112, A
Sequence 8283, Ap	C 23	44.6	6.4	513	9	US-09-785-276A-8283	Sequence 8283, Ap
Sequence 5993, Ap	C 24	44.6	6.4	631	9	US-09-814-353-5993	Sequence 5993, Ap
Sequence 12272, A	C 25	44.6	6.4	631	9	US-09-814-353-12272	Sequence 12272, A
Sequence 5993, Ap	C 26	44.6	6.4	631	9	US-09-814-353A-5993	Sequence 5993, Ap
Sequence 12272, A	C 27	44.6	6.4	631	9	US-09-814-353A-12272	Sequence 12272, A
Sequence 7, Appl1	C 28	44.4	6.3	1799	12	US-10-155-533-7	Sequence 7, Appl1
Sequence 3, Appl1	C 29	44.4	6.3	2017	12	US-10-155-533-3	Sequence 3, Appl1
Sequence 1, Appl1	C 30	44.4	6.3	5314	12	US-10-155-533-1	Sequence 1, Appl1
Sequence 17543, A	C 31	44	6.3	352	9	US-09-814-353-17543	Sequence 17543, A
Sequence 17543, A	C 32	44	6.3	352	9	US-09-814-353A-17543	Sequence 17543, A
Sequence 323, App	C 33	44	6.3	7131	13	US-10-221-613-323	Sequence 323, App
Sequence 1510, Ap	C 34	43.8	6.3	359	9	US-09-789-189-1510	Sequence 1510, Ap
Sequence 7598, Ap	C 35	43.6	6.2	478	9	US-09-785-276A-7598	Sequence 7598, Ap
Sequence 11962, A	C 36	43.4	6.2	491	9	US-09-918-995-11962	Sequence 11962, A
Sequence 914, App	C 37	43.2	6.2	14809	2	PCT-US02-09188-914	Sequence 914, App
Sequence 930, App	C 38	43.2	6.2	14809	2	PCT-US02-09370-930	Sequence 930, App
Sequence 6325, App	C 39	43.2	6.2	14809	12	US-10-105-299-6325	Sequence 6325, App
Sequence 3, Appl1	C 40	43.2	6.2	53332	13	US-10-224-562-3	Sequence 3, Appl1
Sequence 33, Appl1	C 41	43	6.1	396	13	US-10-212-677-33	Sequence 33, Appl1
Sequence 55615, A	C 42	42.8	6.1	512	9	US-09-785-276A-55615	Sequence 55615, A
Sequence 113, App	C 43	42.6	6.1	9760	13	US-10-221-613-113	Sequence 113, App
Sequence 413, App	C 44	42.4	6.1	5241	13	US-10-221-613-413	Sequence 413, App
Sequence 64, Appl1	C 45	42.4	6.1	6298	13	US-10-239-676-64	Sequence 64, Appl1

ALIGNMENTS

RESULT 1

US-10-172-086-13/c

; Sequence 13, Application US/10172086

; GENERAL INFORMATION:

; APPLICANT: Epigenomics AG

; TITLE OF INVENTION: Method and nucleic acids for the differentiation

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/172,086

; CURRENT FILING DATE: 2002-06-13

; NUMBER OF SEQ ID NOS: 116

; SEQ ID NO 13

; LENGTH: 7040

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-172-086-13

Query Match 6.9%; Score 48.4; DB 12; Length 7040;  
Best Local Similarity 52.9%; Pred. No. 2;  
Matches 127; Conservative 0; Mismatches 111; Indels 2; Gaps 1;

Qy	423	AACACAATAAATACACAAAATTTCTCTGAAAGTTCGGAGAGAACTCAATCTCCACA	482
Db	1830	AAAACATATTTTAAATAATATTTAAATAATATCTATCAAACTAAATAAATCG	1771
Qy	483	ATTGGATTTGGACTATATCGATGGATTTCTCTCTCTTAAATTAACATCGCCTATA	542
Db	1770	TTTTCATATTAACATAACATAATTTAAATATTTTCTACTCTTAATTTTACGATTTAA	1711
Qy	543	AAAGAAATGCCACTTGAATAATCAAGAAATCAAAATTAATCAAAATGAATGAAGCC	602
Db	1710	AAATTAATAATCTTACTATATATATATACACATATATATAAACGTAATAAATAATT	1651
Qy	603	ACTAACTCAACAAATCCAGAAATCAGGTGGGAATTAACCTAGAACCAACACAAA	662
Db	1650	A-ACACTAATCTAATCAATACAATATACCTATATCATCCCCCAACAAAAACAAA	1593

RESULT 2

US-10-221-613-204/c  
; Sequence 204, Application US/10221613

; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle  
; FILE REFERENCE: 5013.1004  
; CURRENT APPLICATION NUMBER: US/10/221,613  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: PCT/EP01/02945.

; DE 10013847.00  
; DE 10019058.8  
; DE 10019173.8  
; DE 10032529.7  
; DE 10043826.1  
; PRIOR FILING DATE: 2001-03-15

; 2000-03-15  
; 2000-04-06  
; 2000-04-07  
; 2000-06-30  
; 2000-09-01

; NUMBER OF SEQ ID NOS: 428

; SEQ ID NO 204

; LENGTH: 5501

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-221-613-204

Query Match 6.7%; Score 47.2; DB 13; Length 5501;

Best Local Similarity 46.6%; Pred. No. 3.3;  
Matches 151; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 334 AAGACTTAGCGCACTCCCAATCTCTAATCTCGAATATTATCTTTTCTTACATGGA 393

DB 2933 AATATATAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2874

QY 394 ATAACGTGCAATCTTGATATATTCACAAATTAACAAATTAACAAATTAACAAAT 453

DB 2873 TTAATAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2814

QY 454 AAGTTCGCGAGAACTCAATCTCCCAATTTGGATATTTGGAGTATATCGATGGATT 513

DB 2813 AATATCTTTAAACAAATTTATCTTTAAATACCTTATTAATACTAAATACTATCCAATAA 2754

QY 514 CTTCTCTTTTAAATAACATGCGCTTATAAAGAAATGCGCACTTGAATAATCAAGAT 573

DB 2753 ACAACCTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2694

QY 574 CAAATATTCAAATGAATAATGAAGCCACTAAACTCAACAAATCCAGAAATCAGGTGG 633

DB 2693 AAAAACTTCAAACTAAACATACCATTTAAACAAATAAATAAATAAATAAATAAATAA 2634

QY 634 ATAACCTTAACCTCAGACCAAAAC 657

DB 2633 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2610

RESULT 3

US-09-814-353-6171/c

; Sequence 6171, Application US/09814353

; GENERAL INFORMATION:

; APPLICANT: Lee, John

; APPLICANT: Thompson, Pamela

; APPLICANT: Lillie, James

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

; FILE REFERENCE: MRI-006B

; CURRENT APPLICATION NUMBER: US/09/814,353

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6171

; LENGTH: 639

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 37, 127, 128, 129, 130, 159, 172, 181, 182, 187, 188, 190,

; LOCATION: 217, 218, 219, 228, 229, 230, 232, 233, 234, 239, 240, 253,

; LOCATION: 272, 276, 299, 300, 307, 314, 323, 324, 325, 328, 330, 359,

; LOCATION: 363, 364, 368, 383, 388, 423, 424, 428, 429, 430, 439

; OTHER INFORMATION: n = A,T,C or G

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 444, 447, 452, 458, 463, 464, 470, 472, 474, 481, 482, 484,

; LOCATION: 485, 494, 508, 512, 516, 530, 533, 535, 536, 544, 545, 548,

; LOCATION: 553, 559, 564, 565, 567, 570, 571, 583, 587, 590, 597, 599,

; LOCATION: 608, 610, 613, 614, 619, 623, 624, 634

; OTHER INFORMATION: n = A,T,C or G

; US-09-814-353-6171

Query Match 6.6%; Score 46.2; DB 9; Length 639;

Best Local Similarity 38.4%; Pred. No. 4.1;  
Matches 186; Conservative 0; Mismatches 298; Indels 0; Gaps 0;

QY 175 GAGGTGATATTTTATCTCTGATAGGAACACGCTCTTGAAGCTGTAGCTTATCAATCG 234

DB 534 GNGNGTTTTTTTATTTGTTTNGGANAATAAAGGGGNTTATAAANNNTAATTT 475

QY 235 CTTGTGTTTTCTATACAAACATGTTTGAAGAAGTGAAGCAGCATGCTCTTAACCC 294

DB 474 NTNGNGGGGNNNTAANAACCCCTCTNTTATAAANAANAANNGGNNNTTTTTTT 415

QY 295 TTAATCGCGCTGAGATAAATGACTTTTGGTCTGAGAAAGACTTAGCGACAACCTCC 354

DB 414 TTTTTTTTTTTATTAATAAANAANCCNCCNCCCGGGTTTNAANNNAATNTTTT 355

QY 355 AATCTTCTAATCTGAATATTTATCTTTTCTTACATGGAATCACTTCTTGGATATA 414

DB 354 TTTTTCCTCAAAAAAATTTTNTNGNCCAAAAAATTTTTCCTTTTNNAAA 295

QY 415 TTCCAATTAACAAATAAATCAACAATAATTTCTCTTGAAGTTCGAGAGAACTCAA 474

DB 294 AAAAAAATAATTTTNTGCAAAAAATTTTNTTAAAAAATTTTAAAAAATTTT 235

QY 475 TCTCCACAATTTGGATATTATCGATGGATTTTCTTCTTCTTCTTCTTCTTCTT 534

DB 234 NNNGNNNTTTTNTNGGGGGAATAAATTTTNTTNTGAGNNNAATAA 175

QY 535 CGCCTATAAAGAAATGGCCACTTTGAATTAATCAAGAATCAAAATATTCAAAATGAAATA 594

DB 174 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 115

QY 595 ATGAGCCACTTAACCTCAACAATCCAGATCGAGTGGGAATAACTTAACTCAGAACCA 654

DB 114 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 55

QY 655 AACA 658

DB 54 AAAA 51

```

Qy 475 TCCTCCAAATGGATATATGGAGTATTATCGATGCAATTTCTTCTTTCTTTAAATTAACATG 53A
Db 234 NNNGNNTTTTTTTTTNNNGGGGAAAAAATTTTTTTTTTTTTTTTTNTGAGNNAAAAA 175Y
Qy 535 CGCCTATAAAGAATAATGGCCACTTGAATATCAAGAAATCAAAATATTCAAAATGAATA 59A
Db 174 AANAATATAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 119A
Qy 595 ATGAAGCCACTAAACTCAACAAATCCAGAAATCAGGTGGGAATTAACTTTAATCTCAGAACC 65A
Db 114 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 55
Qy 655 AACA 658
Db 54 AAAA 51

RESULT 5
US-09-814-353A-6171/c
; Sequence 6171, Application US/09814353A
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6171
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 37, 127, 128, 129, 130, 159, 172, 181, 182, 187, 188, 190,
; LOCATION: 217, 218, 219, 220, 232, 233, 234, 239, 240, 253,
; LOCATION: 272, 276, 299, 300, 310, 323, 324, 325, 328, 330, 359,
; LOCATION: 363, 364, 368, 383, 386, 423, 424, 428, 429, 430, 439
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 444, 447, 452, 458, 463, 464, 470, 472, 474, 481, 482, 484,
; LOCATION: 485, 494, 508, 512, 516, 530, 533, 535, 536, 544, 545, 548,
; LOCATION: 553, 559, 564, 565, 567, 570, 571, 583, 587, 590, 597, 599,
; LOCATION: 608, 610, 613, 614, 619, 623, 624, 634
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353A-6171

Query Match 6.6%; Score 46.2; DB 9; Length 639;
Best Local Similarity 38.4%; Pred. No. 4.1;
Matches 186; Conservative 0; Mismatches 298; Indels 0; Gaps

Qy 175 GAGGTGATATTATTCCTTGATAGGAACACGCTCTTGAGCTGTGAAGTATTATTCATCG 234A
Db 534 GNGNGTTTTTTTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 294A
Qy 235 CTCTGTGTTTCTATACAAAACATGTTTTCGAAAAGAATGAAGCACATTTGCTTTAAACCCCT 294A

```







```
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5852
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 74, 82, 152, 171, 177, 190, 195, 198, 207, 210, 215, 216,
; LOCATION: 222, 223, 224, 228, 230, 232, 241, 242, 245, 253, 254, 273,
; LOCATION: 233, 294, 296, 300, 302, 310, 312, 317, 319, 321, 325, 331,
; LOCATION: 333, 335, 336, 346, 348, 352, 369, 373, 375, 383, 388
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 393, 394, 403, 404, 409, 423, 425, 428, 443, 446, 447, 449,
; LOCATION: 451, 455, 465, 468, 469, 472, 474, 475, 478, 481, 488, 490,
; LOCATION: 491, 495, 496, 504, 513, 515, 518, 525, 526
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353A-5852

Query Match          6.6%; Score 46; DB 9; Length 539;
Best Local Similarity 41.7%; Pred. No. 4.3;
Matches 130; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 357 TCTTCTAAATCTCGAATATATCTTTTCCCTTACATGGAATCACTGCATTCTTGGATATATT 416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 TTTTITTTTGGGNANTTTTTTTTNNCNTNCAANAATNTNANATTTNANAATATTN 302

QY 417 CCATTAACACAATAAATCAACAATAATCTCTTGAAGTCCGGAGAAAACCTCAATC 476
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 ANAATNTNTNGAAGAAAAAATTTTTTTTTTTCATCAAAAAANNTTATCCNCCN 242

QY 477 TCCACAATGGATATGGAGTATATCGATGATTTCTCTTCTTAAATCAACATGCG 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 NAAAAATNTNTNGGNNNTAGTCNNTTTNNACNTTTTTTCNCNCTTTTATATTTT 182

QY 537 CCTATAAAGAAATAGCCACTTGAATATCAAGAAATCAAAATATCAAAATGAAATAT 596
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TTTANAAAAANTTTTTTTTCAAAAAAANAATAAAAAAATAAAAAAATAAAAAA 122

QY 597 GAAGCCACTAACTCAACAATCCAGATCAGTGGGGAATACTTAACCTCAGAACCAAA 656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 62

QY 657 CACAAGCCTCA 668
   ||| ||| |||
Db 61 AAAAAAACCCA 50

RESULT 10
US-09-814-353A-12133/c
; Sequence 12133, Application US/09814353A
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
```

```
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12133
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 74, 82, 152, 171, 177, 190, 195, 198, 207, 210, 215, 216,
; LOCATION: 222, 223, 224, 228, 230, 232, 241, 242, 245, 253, 254, 273,
; LOCATION: 233, 294, 296, 300, 302, 310, 312, 317, 319, 321, 325, 331,
; LOCATION: 333, 335, 336, 346, 348, 352, 369, 373, 375, 383, 388
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 393, 394, 403, 404, 409, 423, 425, 428, 443, 446, 447, 449,
; LOCATION: 451, 455, 465, 468, 469, 472, 474, 475, 478, 481, 488, 490,
; LOCATION: 491, 495, 496, 504, 513, 515, 518, 525, 526
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353A-12133

Query Match          6.6%; Score 46; DB 9; Length 539;
Best Local Similarity 41.7%; Pred. No. 4.3;
Matches 130; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 357 TCTTCTAAATCTCGAATATATCTTTTCCCTTACATGGAATCACTGCATTCTTGGATATATT 416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 TTTTITTTTGGGNANTTTTTTTTNNCNTNCAANAATNTNANATTTNANAATATTN 302

QY 417 CCATTAACACAATAAATCAACAATAATCTCTTGAAGTCCGGAGAAAACCTCAATC 476
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 ANAATNTNTNGAAGAAAAAATTTTTTTTTTTCATCAAAAAANNTTATCCNCCN 242

QY 477 TCCACAATGGATATGGAGTATATCGATGATTTCTCTTCTTAAATCAACATGCG 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 NAAAAATNTNTNGGNNNTAGTCNNTTTNNACNTTTTTTCNCNCTTTTATATTTT 182

QY 537 CCTATAAAGAAATAGCCACTTGAATATCAAGAAATCAAAATATCAAAATGAAATAT 596
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TTTANAAAAANTTTTTTTTCAAAAAAANAATAAAAAAATAAAAAAATAAAAAA 122

QY 597 GAAGCCACTAACTCAACAATCCAGATCAGTGGGGAATACTTAACCTCAGAACCAAA 656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 62

QY 657 CACAAGCCTCA 668
   ||| ||| |||
Db 61 AAAAAAACCCA 50

RESULT 11
US-09-814-353-18656/c
; Sequence 18656, Application US/09814353
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
```



```

: LOCATION: 657, 661, 667, 681, 682, 684, 689, 704, 707, 724, 742, 744,
: LOCATION: 751, 755, 758, 773, 774, 781, 782, 783, 784, 786, 798
: OTHER INFORMATION: n = A,T,C or G
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 800, 810, 812, 813, 818, 819, 828, 829, 840, 842, 850, 861,
: LOCATION: 869, 873, 874
: OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5980

```

[illegible]

```

RESULT 14
US-10-221-613-3
; Sequence 3, Application US/10221613
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221.613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 3
; LENGTH: 5680
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-3

```

[illegible]

Db	4506	TTTATATATATGTTGTTTGATTTTTTAACTAAATATTTTTTGGAGAAATAGGGTTTAGGGTTT	4565
Qy	77	ATTTGTTTAAATCCAGATGGATCCGGAATTTTGA AAAACATTTCTAC TAAATTCGAAAT	136
Db	4556	RATTTTTTATAAAGTAGAGGGCGTGGAGAGAAGCGATTTTATTTTTTATTTTTTCGTT	4625
Qy	137	TAAATTTGAGCTTGGCTTGATAAATGGTAGTTGGCTCGGAGGTGATATTTTTATCCTTGA	196
Db	4626	TTAATTTTGTATTTCGATTTAGTTATGCGGTTAAGTTTGGAGGGGACGTTGGTTTTTAGA	4685
Qy	197	TAGGAACACACGCTCTTGAAGCTGTAA GTTTATTCAAATCGCTCTGTTTCTATACAAAAC	256
Db	4686	TATGGTTTCGTTTATTGTAA GTTTATTTTGGTAAGGGAATTTTCCAGTAGAGGTAAC	4745
Qy	257	ATGTTTT	263
Db	4746	CGCAATT	4752

```

RESULT 15
US-09-785-276A-8158
: Sequence 8158, Application US/09785276A
: GENERAL INFORMATION:
: APPLICANT: Schlegel, Robert
: APPLICANT: Endege, Wilson
: APPLICANT: Monahan, John
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
: TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY FOR
: TITLE OF INVENTION: HUMAN PROSTATE CANCER
: FILE REFERENCE: MRI-007B
: CURRENT APPLICATION NUMBER: US/09/785,276A
: CURRENT FILING DATE: 2001-02-16
: PRIOR APPLICATION NUMBER: 60/183,319
: PRIOR FILING DATE: 2000-02-17
: PRIOR APPLICATION NUMBER: 60/189,862
: PRIOR FILING DATE: 2000-03-16
: PRIOR APPLICATION NUMBER: 60/207,454
: PRIOR FILING DATE: 2000-05-25
: PRIOR APPLICATION NUMBER: 60/211,314
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 60/219,007
: PRIOR FILING DATE: 2000-07-18
: PRIOR APPLICATION NUMBER: 60/255,281
: PRIOR FILING DATE: 2000-12-13
: NUMBER OF SEQ ID NOS: 62232
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8158
: LENGTH: 327
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 125, 132, 138, 143, 145, 150, 152, 156, 157, 165, 169, 180,
: LOCATION: 181, 182, 183, 184, 185, 186, 188, 190, 192, 193, 196, 204,
: LOCATION: 208, 212, 220, 223, 226, 228, 229, 232, 234, 239, 240, 244,
: LOCATION: 246, 250, 251, 254, 261, 272, 274, 276, 277, 289, 293
: OTHER INFORMATION: n = A,T,C or G
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 296, 304
: OTHER INFORMATION: n = A,T,C or G
US-09-785-276A-8158

```

	Query Match	6.5%	Score 45.2	DB 9	Length 327
	Best Local Similarity	41.7%	pred. No. 5.9		
	Matches 101	Conservative	0	Mismatches 141	Indels 0
	Gaps				
QY	16	TTTTTTTTTTCTTTTACCTATATTCAAATAGTCTCTTTGATTCAACAGATCGAAAGTCAT	75		
Db	58	TTT	117		
QY	76	TTTTTGTTTTAAATCCAGATGCGCGGAATTTTGAAAACAGATTCTACTAAATTCGAAA	135		

Db 118 TTTTNTTAAATNTAAAAAATANANATTTNTNTATNNCATTACTNAAANAGAAAAA 177  
Qy 136 TTAAATTTGAGCTTGGCTTGATAATGGTAGTTGGCTCGGAGGTGATATTTTATCCCTTG 195  
Db 178 AANNNNNNANANTNNATNTTTATATNTTANATTTATTAANGCNCNANNATNCNAAA 237  
Qy 196 ATAGGAACACGCTTGAAGCTTAAGTTATTCAATCGCTTGCTTTTCTATACAAAAA 255  
Db 238 ANNATANAAAAANNNTANTAAACATNTATTATATTANANTNNGTGTTTTAAANAANAANA 297  
Qy 256 CA 257  
Db 298 TA 299

Search completed: November 1, 2002, 23:31:39  
Job time : 291.458 secs

ALL INFORMATION CONTAINED  
HEREIN IS UNCLASSIFIED  
DATE 07-11-2001 BY 60322 UCBAW

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:16:04 ; Search time 1838.39 Seconds  
(without alignments)  
7331.704 Million cell updates/sec

Title: us-09-881-556-4  
Perfect score: 623  
Sequence: 1 cctcaaatgtgaacatcaa.....aaaaaaaaaaaaaactcgag 623

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 1081749327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_NA\_Main.\*  
1: /cgn2\_6/ptodata/1/pna/PCTUS\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US080\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/1/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq.\*  
18: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq.\*  
19: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq.\*  
20: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq.\*  
21: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq.\*  
22: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq.\*  
23: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq.\*  
24: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq.\*  
25: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq.\*  
26: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq.\*  
27: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq.\*  
28: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq.\*  
29: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq.\*  
30: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq.\*  
31: /cgn2\_6/ptodata/1/pna/US098A\_COMB.seq.\*  
32: /cgn2\_6/ptodata/1/pna/US098B\_COMB.seq.\*  
33: /cgn2\_6/ptodata/1/pna/US098C\_COMB.seq.\*  
34: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq.\*  
35: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq.\*  
36: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq.\*  
37: /cgn2\_6/ptodata/1/pna/US100\_COMB.seq.\*  
38: /cgn2\_6/ptodata/1/pna/US101\_COMB.seq.\*  
39: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq.\*  
40: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq.\*  
41: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq.\*  
42: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq.\*  
43: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq.\*

44: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq.\*  
45: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq.\*  
46: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq.\*  
47: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq.\*  
48: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq.\*  
49: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq.\*  
50: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq.\*  
51: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq.\*  
52: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq.\*  
53: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq.\*  
54: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq.\*  
55: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq.\*  
56: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq.\*  
57: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq.\*  
58: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq.\*  
59: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq.\*  
60: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq.\*  
61: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq.\*  
62: /cgn2\_6/ptodata/1/pna/US6023\_COMB.seq.\*  
63: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq.\*  
64: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq.\*  
65: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq.\*  
66: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq.\*  
67: /cgn2\_6/ptodata/1/pna/US6028\_COMB.seq.\*  
68: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq.\*  
69: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq.\*  
70: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq.\*  
71: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq.\*  
72: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq.\*  
73: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq.\*  
74: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq.\*  
75: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match %	Length	ID	Description
c	1	623	100.0	623	33	US-09-881-556-4 Sequence 4, Appli
c	2	57.8	9.3	787	29	US-09-739-449-2501 Sequence 2501, Ap
c	3	57.8	9.3	787	31	US-09-803-110-2501 Sequence 8029, Ap
c	4	54.2	8.7	543	29	US-09-739-449-8029 Sequence 8029, Ap
c	5	54.2	8.7	543	31	US-09-803-110-8029 Sequence 1859, Ap
c	6	54.2	8.7	1250	29	US-09-739-449-1859 Sequence 1859, Ap
c	7	54.2	8.7	1250	31	US-09-803-110-1859 Sequence 11512, A
c	8	53.2	8.5	562	33	US-09-873-402A-11512 Sequence 11469, A
c	9	53.2	8.5	562	59	US-60-209-830-11469 Sequence 251515, Sequence 15771, A
c	10	53	8.5	458	17	US-09-304-517A-251515 Sequence 63699, A
c	11	53	8.5	458	17	US-09-371-146A-251515 Sequence 251515, Sequence 15771, A
c	12	53	8.5	458	25	US-09-391-630-15771 Sequence 63699, A
c	13	53	8.5	458	25	US-09-654-617-63699 Sequence 251515, Sequence 15771, A
c	14	53	8.5	458	27	US-09-684-016-63699 Sequence 63699, A
c	15	53	8.5	458	36	US-09-985-678-251515 Sequence 251515, Sequence 15771, A
c	16	53	8.5	458	53	US-60-144-084-5658 Sequence 251515, Sequence 15771, A
c	17	52.8	8.5	466	22	US-09-565-306-29077 Sequence 29077, A
c	18	52.6	8.4	398	23	US-09-619-643-16189 Sequence 16189, A
c	19	52.6	8.4	411	25	US-09-654-617-340673 Sequence 340673, Sequence 40673, A
c	20	52.6	8.4	411	27	US-09-684-016-340673 Sequence 340673, Sequence 40673, A
c	21	52.6	8.4	1077	37	US-10-015-127-2817 Sequence 7816, Ap
c	22	52.6	8.4	1077	37	US-10-015-127-2817 Sequence 2817, Ap
c	23	52.2	8.4	1000	33	US-09-881-797-291 Sequence 291, App
c	24	52.2	8.4	1000	37	US-10-021-698-291 Sequence 291, App
c	25	52	8.3	413	24	US-09-637-086A-29308 Sequence 29308, A
c	26	52	8.3	413	24	US-09-637-086B-29308 Sequence 29308, A
c	27	52	8.3	757	29	US-09-739-449-2286 Sequence 2286, Ap
c	28	52	8.3	757	31	US-09-803-110-2286 Sequence 2286, Ap
c	29	51.8	8.3	989	17	US-09-397-761A-2729 Sequence 2729, Ap
c	30	51.8	8.3	989	32	US-09-838-601-2729 Sequence 2729, Ap
c	31	51.6	8.3	664	29	US-09-739-449-739 Sequence 739, App

32 51.6 8.3 664 31 US-09-803-110-739 Sequence 739, App  
c 33 51.4 8.3 637 25 US-09-654-617-429291 Sequence 429291,  
c 34 51.4 8.3 637 26 US-09-669-817A-12145 Sequence 12145, A  
c 35 51.4 8.3 637 27 US-09-684-016-429291 Sequence 429291,  
c 36 51.4 8.3 834 29 US-09-739-449-2256 Sequence 2256, Ap  
c 37 51.4 8.3 834 31 US-09-803-110-2256 Sequence 2256, Ap  
c 38 51.4 8.3 1095 37 US-10-015-127-1938 Sequence 1938, Ap  
c 39 51.2 8.2 530 33 US-09-873-402A-1650 Sequence 1650, Ap  
c 40 51.2 8.2 530 59 US-60-209-830-1650 Sequence 1650, Ap  
c 41 51.2 8.2 592 26 US-09-666-355A-7284 Sequence 7284, Ap  
c 42 51.2 8.2 902 29 US-09-739-449-1929 Sequence 1929, Ap  
c 43 51.2 8.2 902 31 US-09-803-110-1929 Sequence 1929, Ap  
c 44 51 8.2 575 26 US-09-666-355A-12423 Sequence 12423, A  
c 45 51 8.2 612 34 US-09-902-540-1357 Sequence 1357, Ap

ALIGNMENTS

RESULT 1  
US-09-881-556-4  
; Sequence 4, Application US/09881556  
; GENERAL INFORMATION:  
; APPLICANT: Simonsen, J. Neil  
; TITLE OF INVENTION: Cryptosporidium Parvum Antigens, Antibodies Thereto and  
; Diagnostic and Therapeutic Compositions Thereof  
; FILE REFERENCE: 9000-0054  
; CURRENT APPLICATION NUMBER: US/09/881,556  
; CURRENT FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/212,083  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; TYPE: DNA  
; LENGTH: 623  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:C. Parvum AG2  
US-09-881-556-4

Query Match 100.0%; Score 623; DB 33; Length 623;  
Best Local Similarity 100.0%; Pred. No. 3.3e-103; Indels 0; Gaps 0;  
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTCAATGGTGAACATCAAGAACTGTTTCAGAGCAAGCCCTATTACCTCACCAGAAA 60  
Db 1 CCTCAATGGTGAACATCAAGAACTGTTTCAGAGCAAGCCCTATTACCTCACCAGAAA 60  
QY 61 AGCAATCAAGTTCAATCATCTTCTGTGACAGTTCCAGATCTGGATCAGTTCAAAATCT 120  
Db 61 AGCAATCAAGTTCAATCATCTTCTGTGACAGTTCCAGATCTGGATCAGTTCAAAATCT 120  
QY 121 CCTCTGTACTATTCCAGAGACTGGATCAGATCAGATCAGCGCCCTGTGACAAATCC 180  
Db 121 CCTCTGTACTATTCCAGAGACTGGATCAGATCAGATCAGCGCCCTGTGACAAATCC 180  
QY 181 AGAGACTGGATCAGTTCAATCATCTTCTGTACTATATACCAAGACAGATCCAGCTCAG 240  
Db 181 AGAGACTGGATCAGTTCAATCATCTTCTGTACTATATACCAAGACAGATCCAGCTCAG 240  
QY 241 ATCAGACTCTGCTACTTCTCCAGAGAGGATTTGGACTCAGAACGTTTACCAGTCTCT 300  
Db 241 ATCAGACTCTGCTACTTCTCCAGAGAGGATTTGGACTCAGAACGTTTACCAGTCTCT 300  
QY 301 ACAGAACAACTCAAGCCAGCTACATCTTAACCAAGAAATCAAAATCATATATATC 360  
Db 301 ACAGAACAACTCAAGCCAGCTACATCTTAACCAAGAAATCAAAATCATATATATC 360  
QY 361 AGGAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTCACGCA 420  
Db 361 AGGAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTCACGCA 420

QY 421 AGCTGCAGATGTGGAAGCTTATGGGAAAGAGATAAAATGTTGATGTGAGCAAGTAA 480  
Db 421 AGCTGCAGATGTGGAAGCTTATGGGAAAGAGATAAAATGTTGATGTGAGCAAGTAA 480  
QY 481 TCACATAAAATGACATTTATTGAAGTACTTCCGAAGAAATTTAGAACAAATGTAAGTA 540  
Db 481 TCACATAAAATGACATTTATTGAAGTACTTCCGAAGAAATTTAGAACAAATGTAAGTA 540  
QY 541 TCTGCATTCATAAAATGAGCTTATGCTTCCAAATATCTCAAAATGTTCAACTCAAGTAA 600  
Db 541 TCTGCATTCATAAAATGAGCTTATGCTTCCAAATATCTCAAAATGTTCAACTCAAGTAA 600  
QY 601 AAAAAAATAAAATGAGCTTATGCTTCCAAATATCTCAAAATGTTCAACTCAAGTAA 600  
Db 601 AAAAAAATAAAATGAGCTTATGCTTCCAAATATCTCAAAATGTTCAACTCAAGTAA 600

RESULT 2

US-09-739-449-2501/c  
; Sequence 2501, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 09/514,000  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 2501  
; LENGTH: 787  
; TYPE: DNA  
; ORGANISM: Agrobacterium tumefaciens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(787)  
; OTHER INFORMATION: unsure at all n locations  
US-09-739-449-2501

Query Match 9.3%; Score 57.8; DB 29; Length 787;  
Best Local Similarity 50.2%; Pred. No. 0.82;  
Matches 143; Conservative 0; Mismatches 142; Indels 0; Gaps 0;  
QY 333 AACCAAGAAATGAAATCATATAATCAGGAAGCTAATTCAGTTCATTAATACACTAAAT 392  
Db 677 AAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 618  
QY 393 CTTCATAATCAACCAATCTTTCAGCAAGCTGGCAGATGTGGAAGTTATGGGAAAG 452  
Db 617 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 558  
QY 453 GATATAATGGTTGATGGTGAGCAAGTATCACTAATAATGACATATTGACAGTACTTCG 512  
Db 557 TATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 498  
QY 513 AAAGAAATTAAGACAAATAATGTAAGTATCTGCAATTTGATAATATGGCTTAGCCATTC 572  
Db 497 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 438  
QY 573 CAATATCTAATTTGTCACCTCAAGTAAAAAATAAATAAATAAATAAATAAATAAATAA 617  
Db 437 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 393

RESULT 3

US-09-803-110-2501/c  
; Sequence 2501, Application US/09803110  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)D









Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATC 360  
D 301 AAAAAAAAAAAAAAAAAAAAAA 617  
D 403 ATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 87

QY 361 AGGAAGGTAAATCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTCACGCA 420  
D 361 AAAAAAAAAAAAAAAAAAAAAA 617  
D 343 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 284

QY 421 AGCTGCAGATGCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTCACGCA 420  
D 421 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224

QY 481 TCACATAAATGACATTTATGAGATACCTCGAAGAAATAGAAACAAATGTAAGTA 540  
D 481 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224

QY 541 TCTGCATTGATAATATGCGCTTAGCCATTTCCAAATCTAAATGTCACCTCAAGTAA 600  
D 541 AAAAAAAAAAAAAAAAAAAAAA 617  
D 103 TAAAAAATAATATAA 87

RESULT 12  
US-09-391-630-15771/c  
; Sequence 15771, Application US/09391630  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Heck, Gregory R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Shukla, Hridayabhiramjam  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21(15441)B  
; CURRENT APPLICATION NUMBER: US/09/391,630  
; CURRENT FILING DATE: 1999-09-08  
; NUMBER OF SEQ ID NOS: 18889  
; SEQ ID NO 15771  
; LENGTH: 458  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: unsure at all n locations  
US-09-391-630-15771

Query Match 8.5%; Score 53; DB 17; Length 458;  
Best Local Similarity 47.9%; Pred. No. 5.6;  
Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATC 360  
D 301 AAAAAAAAAAAAAAAAAAAAAA 617  
D 403 ATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 87

QY 361 AGGAAGGTAAATCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTCACGCA 420  
D 361 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224

QY 421 AGCTGCAGATGCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTCACGCA 420  
D 421 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224

QY 481 TCACATAAATGACATTTATGAGATACCTCGAAGAAATAGAAACAAATGTAAGTA 540  
D 481 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224

QY 541 TCTGCATTGATAATATGCGCTTAGCCATTTCCAAATCTAAATGTCACCTCAAGTAA 600  
D 541 AAAAAAAAAAAAAAAAAAAAAA 617  
D 103 TAAAAAATAATATAA 87

RESULT 13  
US-09-654-617-63699/c  
; Sequence 63699, Application US/09654617  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; TITLE OF INVENTION: Annotated Plant Genes  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/654,617  
; CURRENT FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 63699  
; LENGTH: 458  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: unsure at all n locations  
US-09-654-617-63699

Query Match 8.5%; Score 53; DB 25; Length 458;  
Best Local Similarity 47.9%; Pred. No. 5.6;  
Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATC 360  
D 301 AAAAAAAAAAAAAAAAAAAAAA 617  
D 403 ATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 87

QY 361 AGGAAGGTAAATCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTCACGCA 420  
D 361 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224

QY 421 AGCTGCAGATGCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTCACGCA 420  
D 421 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224

QY 481 TCACATAAATGACATTTATGAGATACCTCGAAGAAATAGAAACAAATGTAAGTA 540  
D 481 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224

QY 541 TCTGCATTGATAATATGCGCTTAGCCATTTCCAAATCTAAATGTCACCTCAAGTAA 600  
D 541 AAAAAAAAAAAAAAAAAAAAAA 617  
D 103 TAAAAAATAATATAA 87

RESULT 14  
US-09-684-016-63699/c  
; Sequence 63699, Application US/09684016  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; TITLE OF INVENTION: Annotated Plant Genes  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/684,016  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/654,617  
; PRIOR FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 63699  
; LENGTH: 458  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure

LOCATION: (1)..(458)  
OTHER INFORMATION: unsure at all n locations  
US-09-684-016-63699

Query Match 8.5%; Score 53; DB 27; Length 458;  
Best Local Similarity 47.9%; Pred. No. 5.6;  
Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCCTTAACCAAGAAATGAAAAATCATATATC 360  
DB 403 ATATAATAATAAAAAATAGATAAAAAATATAAAAAAATAAAAAATAAAAAATAAT 344  
QY 361 AGGAAGTAAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCAACTTTTCAGCGCA 420  
DB 343 AAAAAAATAATAAAAAATTAATAAAAAAATAAAAAAATAAAAAATAAAAAATAA 284  
QY 421 AGCTGGCAGATGTGGAAGTTTATGGGAAAGAGTAATAAATGGTTGATGGTGAGCAAGTAA 480  
DB 283 AAAAAAATAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATAAAAAATA 224  
QY 481 TCACATAAAATGACATTTTGAAGATACCTCGGAAGAAATTAGAAACAAATGTAAAGTA 540  
DB 223 TAGATAATAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 164  
QY 541 TCTGCATTGATTAATATGGCCTTAGCCATTTCCTCAATATCTAAATTTGTCACCTCAAGTAA 600  
DB 163 AAATTAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 104  
QY 601 AAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 617  
DB 103 TAAAAAATAATATATAA 87

Search completed: November 2, 2002, 00:37:09  
Job time : 1846.39 secs

RESULT 15  
US-09-985-678-251515/c  
Sequence 251515, Application US/09985678  
GENERAL INFORMATION:  
APPLICANT: Cheikh, Nordine  
TITLE OF INVENTION: Annotated plant Genes  
FILE REFERENCE: 16517.255/38-21(15097)F  
CURRENT APPLICATION NUMBER: US/09/985,678  
CURRENT FILING DATE: 2001-11-05  
PRIOR APPLICATION NUMBER: US 09/304,517  
PRIOR FILING DATE: 1999-05-06  
NUMBER OF SEQ ID NOS: 295529  
SEQ ID NO 251515  
LENGTH: 458  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(458)  
OTHER INFORMATION: unsure at all n locations  
US-09-985-678-251515

Query Match 8.5%; Score 53; DB 36; Length 458;  
Best Local Similarity 47.9%; Pred. No. 5.6;  
Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCCTTAACCAAGAAATGAAAAATCATATATC 360  
DB 403 ATATAATAATAAAAAATAGATAAAAAATATAAAAAAATAAAAAATAAAAAATAAT 344  
QY 361 AGGAAGTAAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCAACTTTTCAGCGCA 420  
DB 343 AAAAAAATAATAAAAAATTAATAAAAAAATAAAAAAATAAAAAATAAAAAATAA 284  
QY 421 AGCTGGCAGATGTGGAAGTTTATGGGAAAGAGTAATAAATGGTTGATGGTGAGCAAGTAA 480  
DB 283 AAAAAAATAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 224  
QY 481 TCACATAAAATGACATTTTGAAGATACCTCGGAAGAAATTAGAAACAAATGTAAAGTA 540

Db 223 TAGATAATAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 164  
QY 541 TCTGCATTGATTAATATGGCCTTAGCCATTTCCTCAATATCTAAATTTGTCACCTCAAGTAA 600  
DB 163 AAATTAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 104  
QY 601 AAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 617  
DB 103 TAAAAAATAATATATAA 87

Search completed: November 2, 2002, 00:37:09  
Job time : 1846.39 secs

**THIS PAGE BLANK (USPTO)**











QY 591 ACTCAAGTAAAAA 616  
Db 500 NNAANNNAAAAA 525

RESULT 6

US-10-155-881-3856  
; Sequence 3856, Application US/10155881  
; GENERAL INFORMATION:  
; APPLICANT: Dotson, Stanton B.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Lutfiyya, Linda L.  
; APPLICANT: McIninch, James  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21(15300)J  
; CURRENT APPLICATION NUMBER: US/10/155,881  
; CURRENT FILING DATE: 2002-05-22  
; NUMBER OF SEQ ID NOS: 37595  
; SEQ ID NO 3856  
; LENGTH: 2077  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-155-881-3856

Query Match 7.5%; Score 47; DB 12; Length 2077;  
Best Local Similarity 46.7%; Pred. No. 3.9;  
Matches 149; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 299 CTACAGACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAA 358  
Db 1384 CTAATATAGAGACTAAAGACACCCCTTAAATGTAATGAAGATAACGACAAATGAAGCTCA 1443  
QY 359 TCAGGAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTCACG 418  
Db 1444 GATGATATATGTACATATCTTCGTAAAAA 1503  
QY 419 CAAGCTGGCAGATGTGGAAATGATGGGAAAGGATAAAATGTTGATGGTGAGCAAGT 478  
Db 1504 AAAGGGGAAAAA 1563  
QY 479 AATCACTAAATGACATTTATCAAGATACTTCGAAGAAATTTAGAACCAAAATGTAAAG 538  
Db 1564 AAAAAA 1623  
QY 539 TATCTGATTGATAATATAGCCCTTAGCCCATTTCCAAATATCTAAATTTGTCAACTCAAGT 598  
Db 1624 AAAAAA 1683  
QY 599 AAAAAA 617  
Db 1684 AAAAAA 1702

RESULT 7

US-10-172-086-112/c  
; Sequence 112, Application US/10172086  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Method and nucleic acids for the differentiation  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/172,086  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 112  
; LENGTH: 6944  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-172-086-112

Query Match 7.5%; Score 46.8; DB 12; Length 6944;  
Best Local Similarity 47.9%; Pred. No. 4.8;  
Matches 135; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 333 AACCAAGAAATGAAATCATATAATCAAGAGGTAATTCGAGTTTAAATACACTAAAT 392  
Db 5782 AAAAAA 5723  
QY 393 CTTCCAAATCAACCCCAATCTTTCACGCAAGCTGCGAGATCTGGAAAGTTATGGGAAAG 452  
Db 5722 TTCCAAAAA 5663  
QY 453 GATAAAATGTTGATGGTGAGCAAGTAATCACATTAATAATGACATTAATTTGAAGATACATTCG 512  
Db 5662 AAAAAA 5603  
QY 513 AAAGAAATTTAGAACCAAAATGTAAGTATCTGCAATTCATATAATATGCCCTTAGCCATTTTC 572  
Db 5602 AAAAAA 5543  
QY 573 CAATATCTAAATTTGCAACTCAAGTAATAA 614  
Db 5542 AAAAAA 5501

RESULT 8

US-10-155-881-13051  
; Sequence 13051, Application US/10155881  
; GENERAL INFORMATION:  
; APPLICANT: Dotson, Stanton B.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Lutfiyya, Linda L.  
; APPLICANT: McIninch, James  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21(15300)J  
; CURRENT APPLICATION NUMBER: US/10/155,881  
; CURRENT FILING DATE: 2002-05-22  
; NUMBER OF SEQ ID NOS: 37595  
; SEQ ID NO 13051  
; LENGTH: 2606  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-155-881-13051

Query Match 7.5%; Score 46.6; DB 12; Length 2606;  
Best Local Similarity 46.7%; Pred. No. 4.8;  
Matches 148; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 301 ACAGAACAAATCTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATTC 360  
Db 1767 AAAAAA 1826  
QY 361 AGGAAGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTTCACGCA 420  
Db 1827 AAAAAA 1886  
QY 421 AGCTGCCAGATCTGGAAGTTATGGGCAAGGATAAATCGTTGATCGTGACCAAGTAA 480  
Db 1887 AAAAAA 1946  
QY 481 TCACTAAAAATGACATTTATTTGAAAGTACTTTCGAAAGAAATTTAGAACCAAAATGTAAAGTA 540  
Db 1947 AAAAAA 2006  
QY 541 TCTGCATTGATAATATAGCCCTTAGCCATTTCCAAATATCTAAATTTCTCAACTCAAGTAA 600  
Db 2007 AAAAAA 2066  
QY 601 AAAAAA 617

Db 2067 AAAAAAAAAAAAAAAAAA 2083

RESULT 9

US-10-221-613-93/C

; Sequence 93, Application US/10221613

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle

; FILE REFERENCE: 5013.1004

; CURRENT APPLICATION NUMBER: US/10/221,613

; CURRENT FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: PCT/EP01/02945

; DE 10013847.00

; DE 10019058.8

; DE 10019173.8

; DE 10032529.7

; DE 10043826.1

; PRIOR FILING DATE: 2001-03-15

; 2000-03-15

; 2000-04-06

; 2000-04-07

; 2000-06-30

; 2000-09-01

; NUMBER OF SEQ ID NOS: 428

; SEQ ID NO 93

; LENGTH: 4661

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-221-613-93

Query Match 7.4%; Score 46.4; DB 13; Length 4661;

Best Local Similarity 44.9%; Pred. No. 5.6;

Matches 176; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 224 AACAGGATCCAGCTCAGATCAGACTGCTGCTACTCTCCAGAAAGGATTGGACTCAGAA 283

Db 4134 AAAACAAAAACAAAAAATACATATACATATATCCCAACTCAAAATACATATATCCGATCCAAAC 4075

QY 284 CGTTACCAATCAGCTTCTACAGAACAACTCAAGCCAGCTACATATCTCAACCAAGAAAA 343

Db 4074 CCATCTCTCAATCTCCAAAAAATACCTTTATCTCTACCCCTTATTACAAATACAAAT 4015

QY 344 TGAATCATTAATACAGGAAGTAAATTCGAGTTTAAATACATTAATCTCCAAATCA 403

Db 4014 CAACCTCTCAAAATCAAAATCCTATACATAAAATATTTAAACTAAACCAACCTAAATA 3955

QY 404 ACCCAATCTTCAGCAAGCTGCGAGATGTGGAAGTTATGGGAAAGGATAAATGGT 463

Db 3954 AAACCAACTAAATAAAATCTTATACCTCAAAACACATATAAAACACAAAAAAACCTAA 3895

QY 464 TGATGGTGAAGTAACTACATAAAATGACATTTATTGAAGATACCTTCGAAAGAAATTAG 523

Db 3894 CTCTCCGAAATTAATAAATAAATAAACAATAAATAAATAAATAAATAAATAAATAAATA 3835

QY 524 AAACAAATCTAAGTATCTGCAATGATTAATATGCGCTTTAGCCATTTCCAAATATCTAA 583

Db 3834 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3775

QY 584 ATTGTCACCTCAAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 615

Db 3774 AAAATAAACCAACAAAAACGACCAACCAACAAAA 3743

RESULT 10

US-10-239-676-165/C

; Sequence 165, Application US/10239676

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation

; FILE REFERENCE: 5013.1003

; CURRENT APPLICATION NUMBER: US/10/239,676

; CURRENT FILING DATE: 2002-09-24

; PRIOR APPLICATION NUMBER: PCT/EP01/03968

; DE 10019058.8

; DE 10019173.8

; DE 10032529.7

; DE 10043826.1

; PRIOR FILING DATE: 2001-04-06

; 2000-04-06

; 2000-04-07

; 2000-06-30

; 2000-09-01

; NUMBER OF SEQ ID NOS: 228

; SEQ ID NO 165

; LENGTH: 13606

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-239-676-165

Query Match 7.4%; Score 46.4; DB 13; Length 13606;

Best Local Similarity 47.0%; Pred. No. 6.1;

Matches 143; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 314 AAAGCCAGCTACATATCTTCAACCAAGAAATGAAATCATATAATCAAGGAGTAAATTC 373

Db 11279 AAAATACCTTACAAACCCCTAAAAAACATACCTTAAAAAATAAATAAATAAATAA 11220

QY 374 GAGTTTTAATATACATAAATCTTCCAAATCAACCCAAATCTTTTCACGCAAGCTGGCAGATGT 433

Db 11219 AAAAAAATAAATAATTTCTCCCAACACCCCAAAACTTAAACTAACCCCTAACCAAC 11160

QY 434 GGAAGTTTATGGGAAAGGATAAATGTGTATGTGTGAGCAAGTAAATGTAAGTATCTGCAATTGATA 493

Db 11159 AAAAATAACAAAAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11100

QY 494 CATATTGAGATACATTCGAAAGAAATAGAAACAAATGTAAGTATCTGCAATTGATA 553

Db 11099 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11040

QY 554 ATATGCGCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAAAAAATAAATAA 613

Db 11039 TTATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 10980

QY 614 AAAA 617

Db 10979 AAAA 10976

RESULT 11

US-10-155-881-17885

; Sequence 17885, Application US/10155881

; GENERAL INFORMATION:

; APPLICANT: Dotson, Stanton B.

; APPLICANT: Kovalic, David K.

; APPLICANT: Liu, Jingdong

; APPLICANT: Lutfiyya, Linda L.

; APPLICANT: McIninch, James

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED

; FILE REFERENCE: 38-21(15300)J

; CURRENT APPLICATION NUMBER: US/10/155,881

; CURRENT FILING DATE: 2002-05-22

; NUMBER OF SEQ ID NOS: 37595

; SEQ ID NO 17885

```
; LENGTH: 2729
; TYPE: DNA
; ORGANISM: Glycine max
US-10-155-881-17885

Query Match      7.4%   Score 46.2;   DB 12;   Length 2729;
Best Local Similarity 46.0%;   Pred. No. 5.8;   Mismatches 0;   Gaps 0;
Matches 156;   Conservative 0;

Qy  279  CAGAACGTTACCAATCACTCTCTACAGAACAACTCAAGCCAGCTACATATCCCTAACCAA 338
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2125  CAGCAAGTTATCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2184

Qy  339  GAAATGAAATCATATAATCAGGAGGTAAATTCGAGTTTAAATACACTAAATCTTCCA 398
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2185  AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2244

Qy  399  AATCAACCCCAATCTTTTACGCAAGGTAAATTCGAGTTTAAATACACTAAATCTTCCA 458
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2245  AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2304

Qy  459  ATGCTTGATGCTGAGCAAGTAAATCACTAAATGACATTTATGAAGATACCTCGAAAGAA 518
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2305  AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2364

Qy  519  ATTGAACCAAAATGTAAGTATCTGCATTTGATTAATATGCGCTTAGCCATTTCCAAATA 578
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2365  AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2424

Qy  579  TCTAAATGTCACACTCAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 617
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2425  AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2463

RESULT 12
US-10-198-846-1483
; Sequence 1483, Application US/10198846
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1483
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 46, 52, 75, 78, 80, 85, 91, 100, 102, 107, 109, 114,
; LOCATION: 123, 127, 136, 142, 149, 156, 159, 160, 162, 168, 169, 173,
; LOCATION: 177, 195, 203, 207, 210, 211, 213, 219, 220, 224, 230, 251,
; LOCATION: 261, 288, 277, 309, 327, 333, 337, 339, 355, 359, 364
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 365, 380, 389, 399, 404, 408, 413, 416, 429, 434, 435, 438,
; LOCATION: 440, 458, 473, 479, 490, 511, 519, 520, 522, 525
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1483

Query Match      7.3%   Score 45.6;   DB 13;   Length 525;
Best Local Similarity 39.1%;   Pred. No. 6.5;   Mismatches 0;   Gaps 0;
Matches 156;   Conservative 0;
```

```
Qy  219  CCAGAAACAGGATCCAGCTCAGATCAGACTCTGCTACTCTTCCAGAAAGAGATTGGACT 278
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  38  CCAACCGCGTGGCGGCCGGAAGTAACATTTTNTTNNNNNNNNNNNNNNNNNNNNNNNN 97

Qy  279  CAGAACGTTACCAATCACTCTCTACAGAACAACTCAAGCCAGCTACATATCCCTAACCAA 338
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  98  CANANAGAAANANANANANANANANANANANANANANANANANANANANANANANAN 157

Qy  339  GAAATGAAATCATATAATCAGGAGGTAAATTCGAGTTTAAATACACTAAATCTTCCA 398
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  158  ANNANANANANANANANANANANANANANANANANANANANANANANANANANAN 217

Qy  399  AATCAACCCCAATCTTTTACGCAAGGTAAATTCGAGTTTAAATACACTAAATCTTCCA 458
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  218  ANNANANANANANANANANANANANANANANANANANANANANANANANANANAN 277

Qy  459  ATGCTTGATGCTGAGCAAGTAAATCACTAAATGACATTTATGAAGATACCTCGAAAGAA 518
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  278  AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 337

Qy  519  ATTGAACCAAAATGTAAGTATCTGCATTTGATTAATATGCGCTTAGCCATTTCCAAATA 578
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  338  ANAACCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 397

Qy  579  TCTAAATGTCACACTCAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 617
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  398  TNTAANANANANATATNAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 436

RESULT 13
US-10-221-613-311/c
; Sequence 311, Application US/10221613
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221.613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 311
; LENGTH: 7312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-311

Query Match      7.3%   Score 45.6;   DB 13;   Length 7312;
Best Local Similarity 46.8%;   Pred. No. 8.3;
Matches 144;   Conservative 0;   Mismatches 164;   Indels 0;   Gaps 0;

Qy  310  ACTCAAGCCAGCTACATATCTTACCAAGAAATGAAAAATCAATTAATTAATCAGGAAGCTA 369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2293  ATTAACATCTACTATATATTCTTAACACTATTCCAAAACAAATACAAATTAACAAACA 2234

Qy  370  ATTCGAGTTTTAATACATAAATCTCCCAATCAACCAATCTTTCAGGCAAGCTGGCAG 429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2233  AACCAAAAAAAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2174
```

```
QY 430 ATGTGGAAGTATGCGGGAAGGATAAATGTTGATGGTGAGCAAGTAATCACTAAAA 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2173 ACACAAAAACACCACTTAAAAATAAATAATTCACCTACAATACACCAATATAACAAAT 2114
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 490 ATGACATATTGAGATACCTCGAAGAAATAGAACAAATGTAAGTATCTGCTTG 549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2113 AATAAATTATACTAAACAACCTAAATTAATATAAATAATAAAAAAACTTACTCCAAA 2054
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 550 ATAATATGCGCTTAGCCATTTCCAAATATCTAAATGTCAACTCAAGTAAAAAATAAAA 609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2053 AAATATTTAACTAAACCTCATNAANAATAAANAATCCGACAAATACACCTAAA 1994
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 610 AAAAAAAA 617
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1993 TAAAAAAA 1986
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-10-239-676-211/c
; Sequence 211, Application US/10239676
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 211
; LENGTH: 8996
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (1927..1928, 1943, 1956, 1958, 1981)
US-10-239-676-211

Query Match          7.3%; Score 45.4; DB 13; Length 8996;
Best Local Similarity 48.0%; Pred. No. 9.3;
Matches 159; Conservative 0; Mismatches 171; Indels 1; Gaps 1;

QY 288 ACCAATCACTTCTACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAA 347
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6876 ACAATCTCTTACCCTCATANAACCTTACATCTCTATAAATATATCTATAAAAAA 6817
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 348 ATCATATAAATCAGAGGTAATTCGAGTTTAAATACAC-TAAATCTTCCAAATCAACC 406
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6816 AAAAAACATAAAAAACAAACAATAAATAAATAAATAAACAACAACATTCATAATCATAA 6757
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 407 CAATCTTTCACCGAAGCTGGCAGATGTGGAAGTTATGGGGAAGGATAAATGGTTGA 466
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6756 ACAATAAATAAATAAATAAACACTTACTTTTAAATAAATAAATAAATAAATAAATA 6697
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 467 TGGTGAGCAAGTAATCACTTAAATAATGACATTTATGAGATACCTTCGAAGAAATTTAGAAA 526
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6696 AAAAACTAACATTTTAAATAAATCTTAATAATAAATAAACAATCAATAAATAAATAA 6637
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 527 CAAATGTAAAGTATCTGCGATGATAAATATGCGCTTTAGCCATTTCCAAATCTAAAT 586
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6636 AAAAAATATAAATCCAAACAACAACCTTCTACTTAAAAAATTAATAAACAACATAA 6577
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 587 GTCACTCAAGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 617
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 6576 ATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6546
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 15

```
US-10-239-676-117/c
; Sequence 117, Application US/10239676
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 117
; LENGTH: 11036
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (1927..1928, 1943, 1956, 1958, 1981)
US-10-239-676-117

Query Match          7.2%; Score 45; DB 13; Length 11036;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 135; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 333 AACCAAGAAATGAAATCATATAATCAGGAAGTAATTCAGAGTTTAAATACATAAAT 392
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7784 AACTATAAATAATTCAAAACCTTCTATAACCATATAATAAATAAATAAATAATCAAAAC 7725
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 393 CTTCAAATCAACCAATCTTTTCACGCAAGCTGCAGATGTGGAAGTTATGGGGAAG 452
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7724 TATAAATAAATAACCAATTTATTAATCTATTTTACACAAATAAATAAATTTTTTAAAAA 7665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 453 GATAAATGGTTGATGGTGAGCAAGTAATCACTAAAAATGACATTTATTGAAGATCTTCG 512
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7664 AAAATTAATATTACCACCTTTTAAAAAATCTATAAATAACAAAAACACACATTACAATA 7605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 513 AAGAATAATAGNAACAAAATGTAAAGTATCTGCATTTGTAATATGCGCTTAGCCATTTTC 572
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7604 ACATTAATCTACTAATACCAACTACATATAATATTTTTTAAAAATAAATTAATTTCTAA 7545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 573 CAAATATCTAAATTCCTCAACTCAAGTAAAAAATAAATAAATAAATAAATAAATAA 617
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7544 AAATAATTTAAATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 7500
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 1, 2002, 23:32:35  
Job time : 298.542 secs